

I. INTRODUCTION

The consequences of non-normality, non-additivity, and heterogeneity of error variances on F tests in the analysis of variance have been studied in a number of papers (e.g., Bartlett [1,2], Box and Anderson [3], Cochran [4,5], Curtiss [6], Harter and Lum [7], Nissen and Ottestad [8], Pearson [9], Tukey [10]). It appears that some of the conclusions drawn need further investigation. Also, a statistical procedure designed to determine the scale of measurement required to obtain any of these conditions is not available. Tukey [11] has made some progress in this direction; the statistic designed to determine the appropriate scale of measurement for a set of data to obtain normality is unavailable.

In an attempt to find such a statistic, and in view of the following comment made by Pearson [9] ". . . But in the more extreme cases of non-normal variation, where the standard deviation becomes an unsatisfactory measure of variability, there will always be a danger of accepting the hypothesis when some more efficient test (yet to be devised!) would enable us to pick out a real difference.", it was decided to investigate the minimum coefficient of variation (C.V), a minimum F ratio for Tukey's one degree of freedom for non-additivity, and a maximum F ratios for row and column mean squares. Some analytic results were obtained for the coefficient of variation in relation to positive integer values of k in the transformation X^k . In addition, a Monte Carlo investigation of a series of transformations of the form X^k for $k=\pm 1, \pm 2, \dots, \pm 10$ and for fractional values of $k=\pm .8, \pm .6, \pm .5, \pm .4$, and $\pm .2$ were performed on random normal deviates for various population coefficients of variation (i.e., 5%, 100/15%, 10%, and

* Department of Plant Breeding No. 426, Biometrics Unit No. BU-72.

This work was supported in part by a National Institutes of Health Research Grant RG-5900.

20%). The effect of these various transformations on the C.V. and various F tests was studied.

II. ANALYTICAL RESULTS

Notation: Let X be a random variable normally distributed with the mean μ and variance σ^2 [$N(\mu, \sigma^2)$] where $\mu > 0$ and $\sigma > 0$. Let k be an integer greater than or equal to 1 (i.e., $k=1, 2, 3, \dots$). Let α_k be the k^{th} moment of the random variable X ,

$$\text{i.e., } \alpha_k = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x^k e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

for $k=1, 2, 3, \dots$. α_0 is defined as equal to 1 and α_{-k} is defined as equal to zero.

Lemma: If X is a random variable distributed $N(\mu, \sigma^2)$, $\sigma > 0$ then

$$\alpha_k = \mu\alpha_{k-1} + (k-1)\sigma^2\alpha_{k-2}$$

This can easily be proved directly from the definition of α_k . [13]

The sample coefficient of variation is defined as s/\bar{x} where $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ is the variance and $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ is the mean of the observed sample. The population coefficient of variation can be written as σ/μ (i.e., standard deviation divided by mean for any $f(x)$).

Theorem 1: Let X be a random variable $N(\mu, \sigma^2)$, $\mu > 0$, $\sigma > 0$ and let X^k be the transformation of X for $k=1, 2, 3, \dots$. Then the coefficient of variation of X^k increases as k increases provided $3\mu^4 > \sigma^4$.

Proof: We have to prove

$$\frac{[\alpha_{2k+2} - \alpha_{k+1}^2]^{1/2}}{\alpha_{k+1}} > \frac{[\alpha_{2k} - \alpha_k^2]^{1/2}}{\alpha_k} \quad \text{for } k=1,2,3,\dots$$

Because of the positiveness of the moments and after some simplification it is sufficient to prove that $\alpha_{2k+2} \cdot \alpha_k^2 - \alpha_{2k} \cdot \alpha_{k+1}^2 > 0$ (1)
When k is even the truth of (1) is easily established from the repeated application of the Lemma and it should be noted that in this case the stipulated condition that $3\mu^4 > \sigma^4$ in theorem 1 is not necessary.

When k is odd the proof is a little difficult because the terms in (1) are to be expressed in μ and σ . After some manipulation it can be shown that the terms containing μ and σ of (1) are positive and the term containing σ only is negative. Here only the set condition of the theorem is used to prove (1). Here also it should be noted that for sufficiently large k the condition becomes essentially $\mu > 0$. [13]

Since only the absolute value of the coefficient of variation has meaning in the application the sign can be ignored. In such a case even if $\mu < 0$ (but $\mu \neq 0$) the theorem is true as the magnitude of a moment does not change except the sign and that, too, for odd moments only. If $\mu = 0$ the invalidity of the theorem can easily be seen.

Theorem 2: Let $\log X$ be normally distributed with mean μ and variance σ^2 ($\sigma > 0$). Let X^k be a transformation for all real k. Then the coefficient of variation of X^k increases as $|k|$ increases and coefficient of variation of $X^k = \text{coefficient of variation of } X^{-k}$.

Proof:

$$X^k = e^{k \log X} \quad \text{and} \quad E X^k = E e^{k \log X}$$

The truth of the theorem can easily be seen. [13]*

* Theorem 2 is given on page 11.

III. EMPIRICAL RESULTS

(1) The data and the sampling design: From the book "A Million Random Digits with 100,000 Normal Deviates," by The Rand Corporation [12], pages 21, 55, 107, 159 and 185 of Normal Deviates were randomly selected. On each page there are ten groups of 50 deviates each. The selected pages combined give 50 such groups which we call samples, and each sample represents a two-way classification of five rows and ten columns. To each observation in the sample numbers 5, 10, 15 and 20 were added. This addition results in population means of 5, 10, 15 and 20 with the variance equal to unity in each case. These samples are designated as Category 1. A second set of 50 samples of two rows and ten columns was obtained by taking the first two rows and ten columns of the 50 category 1 samples. This second set of samples was treated in the same manner as the category 1 samples, and was designated as Category 2.

On these two categories a number of transformations $X^{-10}, X^{-9}, X^{-8}, X^{-7}, X^{-6}, X^{-5}, X^{-4}, X^{-3}, X^{-2}, X^{-1}, \log X, X, X^2, X^3, X^4, X^5, X^6, X^7, X^8, X^9, X^{10}$ have been performed and the analyses of variance have been computed. The study of the transformations $X^{-0.8}, X^{-0.6}, X^{-0.5}, X^{-0.4}, X^{-0.2}, X^{0.2}, X^{0.4}, X^{0.5}, X^{0.6}, X^{0.8}$ has also been conducted on the 50 samples of the two categories each with 5 added to each observation of the sample. Log X is said to fit into the above transformations because $(n+1)\int x^n \cdot dx = x^{n+1}$ for all values of n except when $n=-1$ which is $\log X$, i.e., $\int \frac{dx}{x} = \log X$.

The different statistics obtained are shown below.

| Source of Variation | d.f. | F | C.V.% |
|------------------------|----------------|--------------------------|--|
| Rows | $r-1$ | $F_R[(r-1), (r-1)(c-1)]$ | $\frac{\sqrt{\text{Error Mean Square}}}{X_{..}/cr} \times 100$ |
| Columns | $c-1$ | $F_C[(c-1), (r-1)(c-1)]$ | |
| Error | $(r-1)(c-1)$ | | |
| Tukey's Non-Additivity | 1 | $F_T[1, (r-1)(c-1)-1]$ | |
| Residual | $(r-1)(c-1)-1$ | | |

The integer transformations and the averages of the statistics over fifty samples of categories 1 and 2 are given in Tables I and II respectively. Table III contains the fractional transformations and the averages of the statistics of both the categories. (The statistics of individual samples are not given here but may be obtained from the Biometrics Unit, Cornell University, Ithaca, New York.) In order to accommodate the fractional and log transformations, all observations must be positive. Therefore, the minimum number that can be added to each observation in order to have all observations positive is 5. This automatically satisfies the condition of theorem 1.

(2) Coefficient of variation: The coefficient of variation is expressed in percentage.* Figures 1 and 2 represent the graphs of the categories 1 and 2 respectively. The increase in the average coefficient of variation with positive k in the transformation X^k agrees with the theorem 1 as it must. The relatively greater increase in the average C.V. with negative k has not been proved analytically. In fact, no theoretical results were obtained for negative k and $\log X$.

* From here onwards the discussion is on the average values unless otherwise mentioned.

Table I. Averages of F_R , F_C , F_T and C.V.% over 50 samples of category 1. [five rows and ten columns]

| Population means Transformation | 5 | | | | 10 | | | | 15 | | | | 20 | | | |
|------------------------------------|-------|-------|-------------|--------|-------|-------|--------|--------|-------|-------|-------|-------|-------|-------|-------|-------|
| | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% |
| X^{10} | 1.078 | 1.081 | 20.206 | 190.07 | 1.175 | 1.051 | 5.500 | 99.42 | 1.205 | 1.038 | 3.027 | 66.04 | 1.214 | 1.032 | 2.192 | 49.39 |
| X^9 | 1.099 | 1.075 | 15.196 | 170.51 | 1.187 | 1.046 | 4.481 | 88.78 | 1.210 | 1.035 | 2.610 | 59.16 | 1.216 | 1.030 | 1.963 | 44.33 |
| X^8 | 1.121 | 1.069 | 11.230 | 150.78 | 1.198 | 1.042 | 3.635 | 78.36 | 1.214 | 1.032 | 2.252 | 52.38 | 1.216 | 1.028 | 1.763 | 39.32 |
| X^7 | 1.145 | 1.061 | 8.123 | 131.13 | 1.207 | 1.037 | 2.938 | 68.18 | 1.216 | 1.030 | 1.948 | 45.71 | 1.216 | 1.027 | 1.591 | 34.36 |
| X^6 | 1.170 | 1.053 | 5.757 | 111.74 | 1.213 | 1.033 | 2.372 | 58.21 | 1.217 | 1.028 | 1.694 | 39.11 | 1.215 | 1.025 | 1.447 | 29.43 |
| X^5 | 1.193 | 1.045 | 4.005 | 92.78 | 1.217 | 1.030 | 1.921 | 48.41 | 1.215 | 1.026 | 1.487 | 32.58 | 1.212 | 1.024 | 1.328 | 24.53 |
| X^4 | 1.210 | 1.037 | 2.739 | 74.30 | 1.217 | 1.027 | 1.573 | 38.76 | 1.212 | 1.024 | 1.326 | 26.09 | 1.209 | 1.023 | 1.236 | 19.64 |
| X^3 | 1.219 | 1.030 | 1.860 | 56.22 | 1.213 | 1.024 | 1.322 | 29.19 | 1.207 | 1.023 | 1.210 | 19.62 | 1.204 | 1.022 | 1.170 | 14.76 |
| X^2 | 1.214 | 1.025 | 1.314 | 38.28 | 1.204 | 1.022 | 1.168 | 19.60 | 1.200 | 1.021 | 1.140 | 13.13 | 1.198 | 1.021 | 1.130 | 9.86 |
| X | 1.191 | 1.020 | 1.117 | 19.89 | 1.191 | 1.020 | 1.117 | 9.91 | 1.191 | 1.020 | 1.117 | 6.60 | 1.191 | 1.020 | 1.117 | 4.95 |
| $\log X$ | 1.152 | 1.016 | 1.548 | 13.47 | 1.175 | 1.018 | 1.183 | 4.39 | 1.182 | 1.020 | 1.145 | 2.46 | 1.186 | 1.024 | 1.134 | 1.66 |
| X^{-1} | 1.102 | 1.011 | 4.234 | 24.02 | 1.154 | 1.015 | 1.390 | 10.29 | 1.169 | 1.017 | 1.227 | 6.72 | 1.176 | 1.018 | 1.177 | 5.00 |
| X^{-2} | 1.059 | 1.001 | 18.518 | 56.85 | 1.131 | 1.012 | 1.778 | 21.13 | 1.155 | 1.015 | 1.371 | 13.60 | 1.165 | 1.017 | 1.253 | 10.08 |
| X^{-3} | 1.035 | 0.987 | 104.446 | 96.55 | 1.108 | 1.008 | 2.406 | 32.71 | 1.140 | 1.013 | 1.584 | 20.69 | 1.155 | 1.015 | 1.363 | 15.24 |
| X^{-4} | 1.025 | 0.973 | 674.072 | 140.68 | 1.085 | 1.004 | 3.363 | 45.20 | 1.124 | 1.011 | 1.880 | 28.02 | 1.144 | 1.013 | 1.510 | 20.51 |
| X^{-5} | 1.022 | 0.961 | 4698.988 | 188.24 | 1.064 | 0.999 | 4.820 | 58.77 | 1.109 | 1.008 | 2.270 | 35.65 | 1.132 | 1.012 | 1.698 | 25.90 |
| X^{-6} | 1.023 | 0.953 | 34784.568 | 235.97 | 1.046 | 0.993 | 7.020 | 73.49 | 1.093 | 1.005 | 2.774 | 43.61 | 1.121 | 1.010 | 1.933 | 31.42 |
| X^{-7} | 1.024 | 0.947 | 269589.840 | 280.95 | 1.032 | 0.987 | 10.268 | 89.35 | 1.078 | 1.002 | 3.422 | 51.93 | 1.109 | 1.008 | 2.218 | 37.10 |
| X^{-8} | 1.024 | 0.944 | 1582034.800 | 321.62 | 1.021 | 0.982 | 14.852 | 106.25 | 1.065 | 0.999 | 4.262 | 60.65 | 1.098 | 1.006 | 2.561 | 42.95 |
| X^{-9} | 1.024 | 0.942 | *42421.774 | 357.50 | 1.013 | 0.976 | 21.131 | 124.02 | 1.052 | 0.995 | 5.350 | 69.79 | 1.086 | 1.004 | 2.970 | 48.98 |
| X^{-10} | 1.023 | 0.942 | *189815.547 | 388.81 | 1.009 | 0.971 | 29.731 | 142.48 | 1.041 | 0.991 | 6.742 | 79.34 | 1.075 | 1.001 | 3.462 | 55.20 |

*Averages based on 49 samples only. In one case the non-additivity accounted for almost all the error, and so it was excluded.

Table II. Averages of F_R , F_C , F_T and C.V.% over 50 samples of category 2. [First 2 rows and ten columns]

| Population means Transformation | 5 | | | | 10 | | | | 15 | | | | 20 | | | |
|------------------------------------|-------|-------|-----------|--------|-------|-------|--------|--------|-------|-------|--------|-------|-------|-------|-------|-------|
| | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% |
| X^{10} | 1.071 | 1.456 | 32.895 | 174.15 | 1.233 | 1.309 | 5.984 | 93.44 | 1.294 | 1.319 | 3.187 | 62.26 | 1.326 | 1.333 | 2.329 | 46.55 |
| X^9 | 1.107 | 1.399 | 21.938 | 157.81 | 1.255 | 1.309 | 4.810 | 83.62 | 1.309 | 1.325 | 2.754 | 55.78 | 1.337 | 1.338 | 2.103 | 41.77 |
| X^8 | 1.145 | 1.355 | 14.684 | 140.80 | 1.277 | 1.313 | 3.866 | 73.91 | 1.324 | 1.332 | 2.390 | 49.39 | 1.348 | 1.344 | 1.910 | 37.04 |
| X^7 | 1.186 | 1.325 | 9.844 | 123.31 | 1.299 | 1.321 | 3.112 | 64.35 | 1.338 | 1.339 | 2.088 | 43.08 | 1.359 | 1.348 | 1.749 | 32.35 |
| X^6 | 1.228 | 1.310 | 6.593 | 105.59 | 1.320 | 1.330 | 2.518 | 54.93 | 1.353 | 1.346 | 1.843 | 36.84 | 1.369 | 1.353 | 1.619 | 27.69 |
| X^5 | 1.270 | 1.311 | 4.402 | 87.88 | 1.342 | 1.341 | 2.060 | 45.66 | 1.367 | 1.352 | 1.652 | 30.67 | 1.378 | 1.356 | 1.519 | 23.07 |
| X^4 | 1.311 | 1.325 | 2.935 | 70.37 | 1.362 | 1.351 | 1.724 | 36.51 | 1.379 | 1.357 | 1.514 | 24.53 | 1.387 | 1.358 | 1.450 | 18.45 |
| X^3 | 1.351 | 1.346 | 1.992 | 53.13 | 1.381 | 1.358 | 1.505 | 27.45 | 1.390 | 1.359 | 1.430 | 18.43 | 1.394 | 1.360 | 1.411 | 13.86 |
| X^2 | 1.385 | 1.361 | 1.479 | 36.04 | 1.395 | 1.361 | 1.404 | 18.41 | 1.397 | 1.360 | 1.401 | 12.32 | 1.398 | 1.360 | 1.404 | 9.25 |
| X | 1.401 | 1.359 | 1.431 | 18.65 | 1.401 | 1.359 | 1.431 | 9.29 | 1.401 | 1.359 | 1.431 | 6.19 | 1.401 | 1.359 | 1.431 | 4.64 |
| $\log X$ | 1.369 | 1.332 | 2.094 | 12.50 | 1.396 | 1.351 | 1.611 | 4.10 | 1.400 | 1.356 | 1.529 | 2.30 | 1.404 | 1.360 | 1.499 | 1.55 |
| X^{-1} | 1.281 | 1.289 | 4.215 | 21.55 | 1.375 | 1.338 | 1.980 | 9.61 | 1.392 | 1.348 | 1.698 | 6.29 | 1.397 | 1.352 | 1.599 | 4.68 |
| X^{-2} | 1.172 | 1.246 | 10.117 | 47.85 | 1.341 | 1.320 | 2.601 | 19.67 | 1.377 | 1.339 | 1.955 | 12.71 | 1.389 | 1.347 | 1.746 | 9.43 |
| X^{-3} | 1.080 | 1.207 | 27.682 | 79.19 | 1.295 | 1.300 | 3.567 | 30.33 | 1.356 | 1.328 | 2.316 | 19.31 | 1.385 | 1.340 | 1.944 | 14.25 |
| X^{-4} | 1.021 | 1.172 | 85.802 | 113.21 | 1.244 | 1.280 | 5.027 | 41.67 | 1.330 | 1.316 | 2.802 | 26.11 | 1.363 | 1.332 | 2.197 | 19.15 |
| X^{-5} | 0.991 | 1.140 | 296.780 | 146.95 | 1.191 | 1.260 | 7.222 | 53.76 | 1.299 | 1.303 | 3.442 | 33.14 | 1.345 | 1.324 | 2.515 | 24.16 |
| X^{-6} | 0.981 | 1.111 | 1115.845 | 178.61 | 1.141 | 1.241 | 10.500 | 66.59 | 1.266 | 1.290 | 4.276 | 40.44 | 1.325 | 1.314 | 2.908 | 29.29 |
| X^{-7} | 0.983 | 1.087 | 4459.542 | 207.34 | 1.098 | 1.223 | 15.470 | 80.07 | 1.231 | 1.277 | 5.355 | 48.01 | 1.301 | 1.305 | 3.389 | 36.53 |
| X^{-8} | 0.991 | 1.068 | 18575.981 | 232.93 | 1.061 | 1.207 | 23.069 | 94.06 | 1.196 | 1.264 | 6.749 | 55.86 | 1.277 | 1.295 | 3.972 | 39.91 |
| X^{-9} | 1.002 | 1.054 | 82797.723 | 255.39 | 1.031 | 1.191 | 34.816 | 108.35 | 1.163 | 1.251 | 8.549 | 63.98 | 1.251 | 1.285 | 4.678 | 45.43 |
| X^{-10} | 1.013 | 1.042 | 302958.68 | 274.92 | 1.009 | 1.176 | 53.181 | 122.75 | 1.132 | 1.239 | 10.877 | 72.36 | 1.225 | 1.275 | 5.530 | 51.10 |

Table III. Averages of F_R , F_C , F_T and C.V.% over 50 samples of categories 1 and 2.

| Population means Transformation | 5 (Category 2) | | | | 5 (Category 1) | | | |
|------------------------------------|----------------|-------|-------|-------|----------------|-------|-------|-------|
| | F_R | F_C | F_T | C.V.% | F_R | F_C | F_T | C.V.% |
| $X^{0.8}$ | 1.399 | 1.355 | 1.493 | 15.06 | 1.185 | 1.019 | 1.133 | 16.09 |
| $X^{0.6}$ | 1.395 | 1.351 | 1.586 | 11.42 | 1.178 | 1.018 | 1.176 | 12.21 |
| $X^{0.5}$ | 1.392 | 1.348 | 1.645 | 9.57 | 1.174 | 1.018 | 1.209 | 10.24 |
| $X^{0.4}$ | 1.389 | 1.345 | 1.713 | 7.70 | 1.170 | 1.017 | 1.252 | 8.25 |
| $X^{0.2}$ | 1.380 | 1.339 | 1.880 | 3.90 | 1.161 | 1.016 | 1.371 | 4.19 |
| $X^{-0.2}$ | 1.355 | 1.324 | 2.361 | 4.01 | 1.142 | 1.015 | 1.799 | 4.34 |
| $X^{-0.4}$ | 1.339 | 1.316 | 2.693 | 8.16 | 1.132 | 1.014 | 2.151 | 8.86 |
| $X^{-0.5}$ | 1.331 | 1.312 | 2.887 | 10.29 | 1.127 | 1.014 | 2.375 | 11.21 |
| $X^{-0.6}$ | 1.322 | 1.307 | 3.102 | 12.45 | 1.122 | 1.013 | 2.637 | 13.62 |
| $X^{-0.8}$ | 1.302 | 1.298 | 3.603 | 16.91 | 1.112 | 1.012 | 3.308 | 18.65 |

It is interesting to note that the coefficient of variation increases as the transformation gets away from zero, more rapidly on the left (negative integers) than on the right (positive integers), giving a monotonic shape. The shape is approximately convex for the values of k used except when the population coefficient of variation is 20%. As the C.V. of X increases the curves for the left-hand portion of figures 1 and 2 rise more rapidly than right-hand portion. For the right side (positive k) the increase in steepness of the curve is more or less in proportion to the increase in the C.V. of X . For smaller C.V. the difference between the slopes on the two sides is smaller than for large C.V. Nevertheless, the C.V. of the transformations for negative k is greater than the corresponding transformations for positive k . Considering the integer transformations but excluding the zero transformation one can note that the C.V. of the log transformation is minimum by virtue of its position in between a coefficient of variation of zero and the coefficient of variation for $k=1$. Actually the transformation $\log X$ cannot be considered as the transformation for $k=0$ because the coefficient of variation of the latter equals zero and the coefficient of variation of $\log X$ is not zero. $\log X$ had the minimum C.V. irrespective of the sample chosen for $|k| \geq 1$. As long as the conditions in theorem 1 hold the increase in C.V. when the transformation gets away from zero does not depend on the random sample selected unless all sample values equal a constant; in none of the analyses of variance on the 400 samples was there an exception to this.

With respect to one or more of the properties mentioned above, the number of cases not satisfying them does depend primarily on the C.V. of X and also on the size of the sample chosen. As the C.V. of X decreases the

the discrepant cases increase and at the same time the discrepancy becomes negligible. Hence as the C.V. of X approaches zero, the curve of the C.V. approaches the abscissa. The difference between the C.V.'s of transformation for positive k and the corresponding negative k increases as the mod of the transformation increases. In any case, the C.V. of the transformation whose mod is greater than 1 is always greater than the C.V. of X .

The following discussion with respect to samples is centered around the transformations for $k=\pm 1$ which is important in view of the conclusions drawn later on. The following theorem is needed for this discussion. Following the same notation in section II it can be proved by direct evaluation that:

Theorem 2: $\alpha_{2k} \leq \alpha_k^2 \left[1+k^2 \cdot \frac{\sigma^2}{\mu^2} \right]$ for $k=1,2,3,4$ and 5 with no restrictions on μ and σ except that $\sigma > 0$. It is a strict inequality for $k > 1$ and equality for $k=1$. [13]

In most cases the C.V. of X is less than C.V. of X^{-1} . For the 400 samples, the C.V. of X is definitely less than the C.V. of X^{-1} by a suitable margin in 64.5% of the cases. In the remaining cases it may be considered that they are equal for the following reasons:

a) Table IV gives an idea of the number of cases and the margin of the difference existing between the C.V. of X and C.V. of X^{-1} when the latter is

Table IV. Range and % of cases where C.V. $X^{-1} < \text{C.V. } X$

| Number added Cate- gories | 5, 20% [†] | | 10, 10% [†] | | 15, 6.7% [†] | | 20, 5% [†] | |
|------------------------------------|---------------------|------------|----------------------|------------|-----------------------|------------|---------------------|------------|
| | Range | % of cases | Range | % of cases | Range | % of cases | Range | % of cases |
| 1 | 0.22-0.37 | 4 | 0.005-0.26 | 20 | 0.005-0.153 | 30 | 0.001-0.097 | 32 |
| 2 | 0.15-1.55 | 18 | 0.04-0.48 | 38 | 0.00-0.24 | 42 | 0.00-0.14 | 44 |

[†] population C.V. in percentage

less than the former. For all combinations of sample size and the C.V. of X the highest percentage of cases where the C.V. of X^{-1} was less than the C.V. of X was 44%. The size of the difference, however, was small. As the C.V. of X increases the percentage becomes smaller. The range of the difference is negligible except in the second row first column where there are 8% of cases where the difference > 1 but < 2 .

b) For the four samples of category 2, selected from a population with a coefficient of variation 20% and where the difference (C.V. X - C.V. X^{-1}) is greater than 1 but less than 2 the C.V. was calculated from a two-way classification (all the C.V.'s are calculated from two-way classifications) and corresponding C.V. was calculated from the variation among the 20 observations in the sample. The differences in two samples were considerably reduced, in one sample it turned out to be negative (i.e., C.V. $X < \text{C.V. } X^{-1}$) and in one sample the difference did not change appreciably.

c) Except for one case in row 2 column 1 all other cases occurred where the inequality in theorem 2 has been violated.

For all the other transformations for k , whether they are fractional or integer, results similar to those in Table IV were obtained. (c) also holds for all of these analyses.

(3) Row F (F_R) and column F (F_C): Figures 3 and 4 and 5 and 6 show the behaviour of F_R and F_C respectively, with respect to the transformations for changing coefficients of variation of X . In individual samples the behaviour of F_R and F_C is not as shown in the figures. It depends upon the particular observations in the sample.

Looking at figures 3, 5 and 6 one finds that normality need not give maximum F values nor that a transformation to attain normality necessarily

gives an increase in F . Although the transformations left of 1 always have a lower F , this is not true with respect to individual samples where sometimes these transformations have a higher F than the F for $k=1$. The magnitude of the differences in the F values as the transformation departs from $k=1$ depends upon the sample chosen. The value of k in the transformation from where the F takes either a downward or an upward turn depends on the individual sample for X for $k=1$. This indicates that F oscillates between certain limits. The range of the limits depends on the coefficient of variation of X . This aspect will be discussed later.

Not all the transformations to the right of $k=1$ have higher F 's than the F for $k=1$. At some value of k in the transformation, F reaches a maximum. This point depends on the degrees of freedom of the numerator as well as the denominator, but to a greater extent on the degrees of freedom of the denominator. The higher the d.f. of the denominator the higher the value of k in the transformation required to reach a maximum. Figures 3, 4, 5 and 6 substantiate this statement. The effect of the d.f. of the denominator is more clearly seen by observing figures 5 and 6 where the numerator d.f. are the same (i.e. 9). The F values of the transformations for k negative are less than the F values of the corresponding transformations for k positive.

From figure 4 one is tempted to come to a conclusion that F_R attains a maximum almost always when the normality condition is satisfied and the numerator has one degree of freedom. But when individual samples are considered one finds that such a conclusion is erroneous. For transformations with integer values of k on the 50 samples from category 2 the number of samples where F_R attains a maximum for $k=1$ are 3, 2, 1 and zero respectively for the population means 5, 10, 15 and 20. If fractional transformations are

also considered then the number of cases is 2 for $\mu=5$. However, on the average F_R attains a maximum when the transformation is $k=1$ and the numerator d.f. is one. This shows that no transformation uniquely gives a maximum F_R . Secondly, even though the maximum F_R may be attained for transformations with k other than one the proportional increase, i.e. $(F_R \text{ for } k \neq 1)/(F_R \text{ for } k=1) > 1$, is less than the proportional decrease, i.e., $(F_R \text{ for } k=1)/(F_R \text{ for } k \neq 1) > 1$, and the above two ratios increase as the mod of the transformation becomes larger. Even in this case F_R oscillates between certain limits.

From figures 3, 4, 5 and 6 it may be seen that the normality condition in the Analysis of Variance cannot be taken easily as is common among statisticians (see, for example, Harter and Lum [7]). The following results of some individual samples (Table V) show some serious effects of non-normality. The seriousness of non-normality depends upon the coefficient of variation of the sample chosen and the transformation required to attain normality.

(4) Tukey's non-additivity F (F_T): Figures 7 and 8 show the behaviour of F_T with respect to the transformations for different co-efficients of variation. In both the figures in the case of $\mu=5$, F_T has not been graphed for some of the transformations with k negative because of large values which are given in tables I and II. In all cases in figure 7 and for $\mu=5$ in figure 8 the F_T is minimum when normality is satisfied. For other population means in figure 8 F_T reaches its minimum at the transformation $k=2$ which is not clearly visible because of compressed scale. After attaining the value of 1.401, F_T begins to rise, which may be taken as an indication that the transformations other than $k=1$ may in general have higher F_T than the transformation $k=1$. Even though, in general, the graph of F_T of individual samples

Table V. F_R values of four individual samples and transformations

| Transformation | $F_R [1,9]$ | | | |
|----------------|-------------|--------|--------|---------|
| X^{10} | 0.232 | 0.351 | 6.924 | 0.431 |
| X^9 | 0.450 | 0.312 | 7.498 | 0.399 |
| X^8 | 0.783 | 0.268 | 7.984 | 0.363 |
| X^7 | 1.267 | 0.216 | 8.348 | 0.321 |
| X^6 | 1.950 | 0.157 | 8.573 | 0.272 |
| X^5 | 2.893 | 0.093 | 8.656* | 0.213 |
| X^4 | 4.159 | 0.033 | 8.607 | 0.147 |
| X^3 | 5.789 | 0.0007 | 8.446 | 0.079 |
| X^2 | 7.680 | 0.024 | 8.192 | 0.023 |
| $X[N(5,1)]$ | 9.371 | 0.117 | 7.862 | 0.00006 |
| $X^{0.8}$ | 9.610 | 0.142 | 7.788 | 0.002 |
| $X^{0.6}$ | 9.797 | 0.170 | 7.712 | 0.006 |
| $X^{0.5}$ | 9.869 | 0.185 | 7.673 | 0.009 |
| $X^{0.4}$ | 9.926 | 0.200 | 7.634 | 0.013 |
| $X^{0.2}$ | 9.990* | 0.231 | 7.554 | 0.022 |
| $\log X$ | 9.988 | 0.262 | 7.472 | 0.035 |
| $X^{-0.2}$ | 9.918 | 0.295 | 7.387 | 0.050 |
| $X^{-0.4}$ | 9.781 | 0.329 | 7.302 | 0.069 |
| $X^{-0.5}$ | 9.690 | 0.345 | 7.258 | 0.079 |
| $X^{-0.6}$ | 9.583 | 0.362 | 7.214 | 0.090 |
| $X^{-0.8}$ | 9.330 | 0.396 | 7.126 | 0.114 |
| X^{-1} | 9.030 | 0.430 | 7.035 | 0.141 |
| X^{-2} | 7.145 | 0.594 | 6.568 | 0.313 |
| X^{-3} | 5.303 | 0.743 | 6.084 | 0.527 |
| X^{-4} | 3.918 | 0.874 | 5.600 | 0.754 |
| X^{-5} | 2.979 | 0.984 | 5.132 | 0.971 |
| X^{-6} | 2.362 | 1.072 | 4.693 | 1.164 |
| X^{-7} | 1.955 | 1.140 | 4.290 | 1.328 |
| X^{-8} | 1.683 | 1.188 | 3.928 | 1.464 |
| X^{-9} | 1.498 | 1.219 | 3.607 | 1.575 |
| X^{-10} | 1.369 | 1.235* | 3.326 | 1.665* |

* Maximum value

has a U-shape, the minimum is not always at the transformation for $k=1$. Table VI gives an idea of the number of times a transformation has the lowest F_T among the 50 samples.

Table VI gives the number of times that a transformation has the lowest F_T and significant F_T at 5% and 1% levels and the total, in both the categories for the different population means. The number of times F_T significant at 5% level does not include the number of times F_T significant at 1% level. Except in the case of $\log X$ in category 1 for $\mu=20$ the total number of times that the F_T is significant with respect to the transformation $k=1$ is never greater than the totals in other transformations for all population means of the two categories. It is interesting to note that in category 1 the total does not decrease as the transformation departs from $k=1$; the same pattern does exist in category 2 also except for the two transformations $k=2$ and 3. As the population mean increased the total either decreased or stayed stationary except for transformations $k=2$ and 3 in category 2. In category 2 for the transformation $k=2$ the total increased when $\mu=10$ and stayed stationary for the rest of the population means; for the transformation $k=3$ the total was the same for $\mu=5, 10$ and 15 , but increased when $\mu=20$.

Not all transformations have significant F_T in the samples where the transformation $k=1$ has a significant F_T thereby confirming that the non-normality has its effect on the type I error. So a transformation to attain additivity without satisfying normality is of no use for hypothesis testing. The effect of increase in the d.f. of the denominator depends on the coefficient of variation of the sample. But for the transformations greater than $k=3$ the situation deteriorates.

Table VI. Number of times a transformation has lowest and significant F_T .

| Population means Transformation | Category 1 F _T (1,35) | | | | | | | | | | Category 2 F _T (1,8) | | | | | | | | | | | | | | | | | | | | | |
|------------------------------------|-------------------------------------|----|-------|----|----|----|-------|----|----|----|------------------------------------|----|----|----|-------|----|----|----|-------|----|----|----|-------|----|----|----|-------|----|---|---|----|----|
| | 5 | | | 10 | | | 15 | | | 20 | | | 5 | | | 10 | | | 15 | | | 20 | | | | | | | | | | |
| | A | | | B | A | | | B | A | | | B | A | | | B | A | | | B | A | | | B | | | | | | | | |
| | 5% | 1% | Total | | 5% | 1% | Total | | 5% | 1% | Total | | 5% | 1% | Total | | 5% | 1% | Total | | 5% | 1% | Total | | 5% | 1% | Total | | | | | |
| X ¹⁰ | 4 | 30 | 34 | 1 | 9 | 13 | 22 | 4 | 11 | 4 | 15 | 5 | 3 | 3 | 6 | 9 | 5 | 22 | 27 | 2 | 6 | 7 | 13 | 7 | 6 | 2 | 8 | 10 | 2 | 2 | 4 | 12 |
| X ⁹ | 5 | 26 | 31 | | 7 | 12 | 19 | 1 | 8 | 4 | 12 | 1 | 3 | 2 | 5 | 3 | 9 | 17 | 26 | | 8 | 5 | 13 | | 4 | 2 | 6 | | 2 | 2 | 4 | |
| X ⁸ | 8 | 23 | 31 | | 10 | 6 | 16 | | 5 | 4 | 9 | 3 | 3 | 2 | 5 | | 10 | 14 | 24 | | 8 | 4 | 12 | 1 | 2 | 2 | 4 | 2 | 1 | 2 | 3 | |
| X ⁷ | 12 | 16 | 28 | 2 | 9 | 4 | 13 | | 3 | 2 | 5 | 3 | 3 | 2 | 5 | | 10 | 11 | 21 | 1 | 6 | 2 | 8 | 2 | 2 | 2 | 4 | | 1 | 2 | 3 | 1 |
| X ⁶ | 9 | 14 | 23 | 1 | 6 | 4 | 10 | 3 | 3 | 2 | 5 | | 3 | 2 | 5 | 4 | 5 | 9 | 14 | 3 | 3 | 2 | 5 | 2 | 1 | 2 | 3 | 1 | 2 | 1 | 3 | 2 |
| X ⁵ | 8 | 9 | 17 | 1 | 3 | 2 | 5 | 5 | 3 | 2 | 5 | 2 | 2 | 1 | 3 | 1 | 9 | 4 | 13 | 2 | 2 | 2 | 4 | | 1 | 2 | 3 | 2 | 2 | 1 | 3 | 5 |
| X ⁴ | 8 | 4 | 12 | 1 | 3 | 2 | 5 | 1 | 2 | 1 | 3 | 2 | 1 | 1 | 2 | 1 | 6 | 2 | 8 | 4 | 1 | 2 | 3 | 3 | 2 | 1 | 3 | 5 | 2 | 1 | 3 | 1 |
| X ³ | 3 | 2 | 5 | 7 | 2 | 1 | 3 | 4 | 1 | 1 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 3 | 1 | 2 | 1 | 3 | 5 | 2 | 1 | 3 | 3 | 3 | 1 | 4 | 3 |
| X ² | 2 | 1 | 3 | 6 | 1 | 1 | 2 | 3 | 1 | 1 | 2 | 3 | 1 | 1 | 2 | 3 | 2 | 1 | 3 | 10 | 3 | 1 | 4 | 4 | 3 | 1 | 4 | 2 | 3 | 1 | 4 | 1 |
| X | 1 | 1 | 2 | 11 | 1 | 1 | 2 | 8 | 1 | 1 | 2 | 6 | 1 | 1 | 2 | 3 | 3 | 0 | 3 | 9 | 3 | 0 | 3 | 6 | 3 | 0 | 3 | 4 | 3 | 0 | 3 | 4 |
| log X | 4 | 1 | 5 | 5 | 3 | 0 | 3 | 2 | 2 | 0 | 2 | 2 | 1 | 0 | 1 | 4 | 5 | 1 | 6 | 7 | 2 | 1 | 3 | 3 | 3 | 0 | 3 | 3 | 3 | 0 | 3 | 2 |
| X ⁻¹ | 6 | 4 | 10 | 6 | 4 | 0 | 4 | 1 | 4 | 0 | 4 | 1 | 3 | 0 | 3 | 2 | 7 | 4 | 11 | 7 | 4 | 1 | 5 | 2 | 4 | 1 | 5 | 2 | 2 | 1 | 3 | 2 |
| X ⁻² | 10 | 10 | 20 | 3 | 4 | 1 | 5 | 4 | 4 | 0 | 4 | 1 | 4 | 0 | 4 | | 9 | 8 | 17 | 1 | 6 | 1 | 7 | 5 | 4 | 1 | 5 | 2 | 4 | 1 | 5 | 1 |
| X ⁻³ | 8 | 20 | 28 | 4 | 5 | 3 | 8 | 3 | 5 | 0 | 5 | 2 | 4 | 0 | 4 | 1 | 12 | 13 | 25 | 1 | 3 | 4 | 7 | 3 | 6 | 1 | 7 | 1 | 4 | 1 | 5 | 1 |
| X ⁻⁴ | 9 | 27 | 36 | 1 | 11 | 4 | 15 | 2 | 2 | 3 | 5 | 2 | 4 | 0 | 4 | | 10 | 22 | 32 | | 7 | 5 | 12 | 4 | 4 | 3 | 7 | 4 | 6 | 1 | 7 | 1 |
| X ⁻⁵ | 5 | 35 | 40 | | 8 | 8 | 16 | 1 | 5 | 3 | 8 | 2 | 4 | 1 | 5 | 3 | 9 | 26 | 35 | 1 | 8 | 7 | 15 | | 4 | 3 | 7 | 2 | 6 | 1 | 7 | 2 |
| X ⁻⁶ | 8 | 37 | 45 | 1 | 7 | 13 | 20 | 2 | 6 | 4 | 10 | 1 | 2 | 3 | 5 | 1 | 5 | 31 | 36 | | 8 | 10 | 18 | | 5 | 5 | 10 | 4 | 4 | 3 | 7 | 3 |
| X ⁻⁷ | 5 | 42 | 47 | | 5 | 18 | 23 | 2 | 11 | 4 | 15 | 3 | 5 | 3 | 8 | 2 | 2 | 35 | 37 | | 9 | 13 | 22 | | 6 | 5 | 11 | 1 | 4 | 3 | 7 | 2 |
| X ⁻⁸ | 2 | 45 | 47 | | 7 | 20 | 27 | 1 | 10 | 6 | 16 | 1 | 4 | 4 | 8 | 1 | 1 | 37 | 38 | 1 | 12 | 13 | 25 | | 8 | 7 | 15 | | 4 | 4 | 8 | 2 |
| X ⁻⁹ | 0 | 47 | 47 | | 11 | 22 | 33 | 1 | 9 | 9 | 18 | 1 | 9 | 4 | 13 | 1 | 2 | 37 | 39 | | 9 | 18 | 27 | | 9 | 8 | 17 | | 6 | 5 | 11 | 3 |
| X ⁻¹⁰ | 0 | 47 | 47 | | 11 | 25 | 36 | 2 | 7 | 13 | 20 | 7 | 11 | 4 | 15 | 9 | 3 | 38 | 41 | | 9 | 21 | 30 | 3 | 8 | 10 | 18 | 2 | 7 | 5 | 12 | 2 |
| Total | | | | 50 | | | | 50 | | | | 50 | | | | 50 | | | | 50 | | | | 50 | | | | 50 | | | | 50 |

A: No. of times F_T significant

B: No. of times F_T is minimum

It is difficult to explain figures 7 and 8 unless one studies each transformation and each sample, but one may generally say that the number of times a transformation has lowest F_T is often counteracted by the total number of times the transformation attains a significant F_T . In individual samples the lowest F_T does not insure the highest F_R or F_C . Simply increasing the d.f. of the denominator is of no value unless the extent of non-normality is known. In both the categories on the average the F_T of the transformations with positive k (integers as well as fractional) is less than the corresponding transformations with negative k .*

(5) Relation of the coefficient of variation and F 's: There is a considerable effect of the change in the coefficient of variation on the F 's, especially when the value of k in the transformation becomes large. A substantial increase in the coefficient of variation might distort the conclusions even though there is slight non-normality. The violent effect of the change in the coefficient of variation is more clearly seen in F_T and especially in the transformations with k equal to negative integers. The disadvantage of having smaller d.f. in the denominator may be offset by decreasing the coefficient of variation.

Going back to figures 3, 4, 5, 6, 7 and 8 one can observe that as the coefficient of variation decreases the F 's of different transformations approach the values of the transformation $k=1$. This is also true in individual samples. The maxima and minima in figures 3, 4, 5 and 6 shift in position farther from the origin as the coefficient of variation decreases. The

* In none of the samples does the transformation $k=1$ have the highest F_T in category 2 and only in one of the samples does it have the highest F_T in category 1, with a value 0.826.

range within which F_R and F_C fluctuate decreases as the C.V. becomes smaller. The log transformation with low coefficients of variation has a tendency to cause irregularities in the curves in figures 4, 5 and 6.

With respect to individual samples and for positive k the position of the minimum F_T shifts farther to the right for some samples and stays stationary for other samples as the coefficient of variation decreased. The reverse was true for negative k . This was true for all but three samples in category 1 and one sample in category 2. In figure 8 the minimum F_T is not attained for $k=1$ when the population means are 10, 15 and 20. The transformations $k=2$ and 3 have a value F_T less than the transformation $k=1$. This may partly be explained by the irregular behaviour of these transformations in specific samples where as in other samples the increasing F_T was accompanied either by an increase or a decrease in the C.V. The change in coefficient of variation has no effect on the F values for the transformation $k=1$.

IV. CONCLUSIONS

An experimental statistician usually draws a set of conclusions from an empirical investigation. Here the conclusions are drawn in the form of conjectures which may be proved or disproved. The following conjectures are under the conditions mentioned in theorem 1 in section II.

- a) The C.V. of X is less than the C.V. of X^k where k is a real number and $|k| > 1$.
- b) C.V. of $X^k \leq$ C.V. of X^{-k} for all real k .
- c) $\alpha_{2k} \leq \alpha_k^2 [1+k^2 (\frac{\sigma}{\mu} + c)^2]$ for all positive integer values of k and for all μ 's, and " c " a constant such that $|c| < 1$.
- d) Log X has a smaller coefficient of variation than any X^k for $|k| \geq 1$.
- e) The maximum F value is a function of k , μ and σ^2 .
- f) For all the transformations of X^k (k any real number not equal to zero) minimum F_7 is attained when $k=1$.
- g) In analyses of variance on X^k for $k \neq 1$ there is a positive relation between the discrepancy in the Type I error and the coefficient of variation.
- h) For a given population coefficient of variation there is a positive relation between the discrepancy in the Type I error and the increase in the denominator d.f. and non-normality.
- i) The greater the C.V. the greater the requirement of normality in hypothesis testing.

V. REMARKS

The validity of testing of hypothesis and of constructing confidence intervals in the analysis of variance are based on the fulfillment of the assumptions of additivity, normality and homogeneity of variances. In practice an experimenter may encounter situations in which one or more of these assumptions are violated. In an attempt to find a statistic which can detect normality and indicate the value of k resulting in normality a number of transformations of the form X^k have been studied on a number of random samples drawn from normal populations with population coefficients of variation 20%, 10%, 100/15% and 5% (i.e., with population means 5, 10, 15 and 20 and variance 1 in each). From this study it has been concluded that: (1) the maximum row and column F 's are unreliable as statistics to indicate normality and/or additivity. Also, minimum F_7 is an unreliable statistic to indicate normality; (2) normality is essential in order to draw valid conclusions in hypothesis testing; (3) the fulfillment of normality also achieves the condition of additivity to a considerable extent.

The relation between various coefficients of variation and various F 's for the integer transformations for $|k| \geq 1$ leads one to conclude that use of X^k giving minimum coefficient of variation removes non-normality and non-additivity. Therefore, the minimum coefficient of variation may be recommended as a statistic to determine the scale of measurement to be analyzed. To determine the appropriate scale the following procedure may be adopted: Conduct a series of transformations of the form X^k where $|k| \geq 1$ and pick out the transformation resulting in the minimum coefficient of variation. This X^k may then be validly used in hypothesis testing or interval estimation,

provided that the transformation is of the prescribed form and that the assumption of homogeneity of variances is not violated. After getting the value of k as mentioned above, if one finds inequality of variances a transformation may be tried within the range (i.e., between zero and the value of k) to stabilize the variances.

If the slopes of the curves of coefficient of variation of positive and negative values (not necessarily integers) of k are equal then log transformation gives normality. [13]

VI. SUMMARY

A number of analytical results were obtained for the statistics known as the coefficient of variation, i.e., standard deviation/mean resulting in three theorems; these are:

Theorem 1: Let X be a random variable normally distributed with mean $\mu > 0$ and variance $\sigma^2 (\sigma > 0)$. Let X^k be a transformation for $k=1,2,3,\dots$. Then the coefficient of variation of X^k increases as k increases provided $3\mu^4 > \sigma^4$.

Theorem 2: Let X be a random variable normally distributed with mean μ and variance $\sigma^2 (\sigma > 0)$. Let α_k be the k^{th} moment. Then $\alpha_{2k} \leq \alpha_k^2 \left[1 + k^2 \frac{\sigma^2}{\mu^2} \right]$ for $k=1,2,3,4$ and 5 with equality when $k=1$ only.

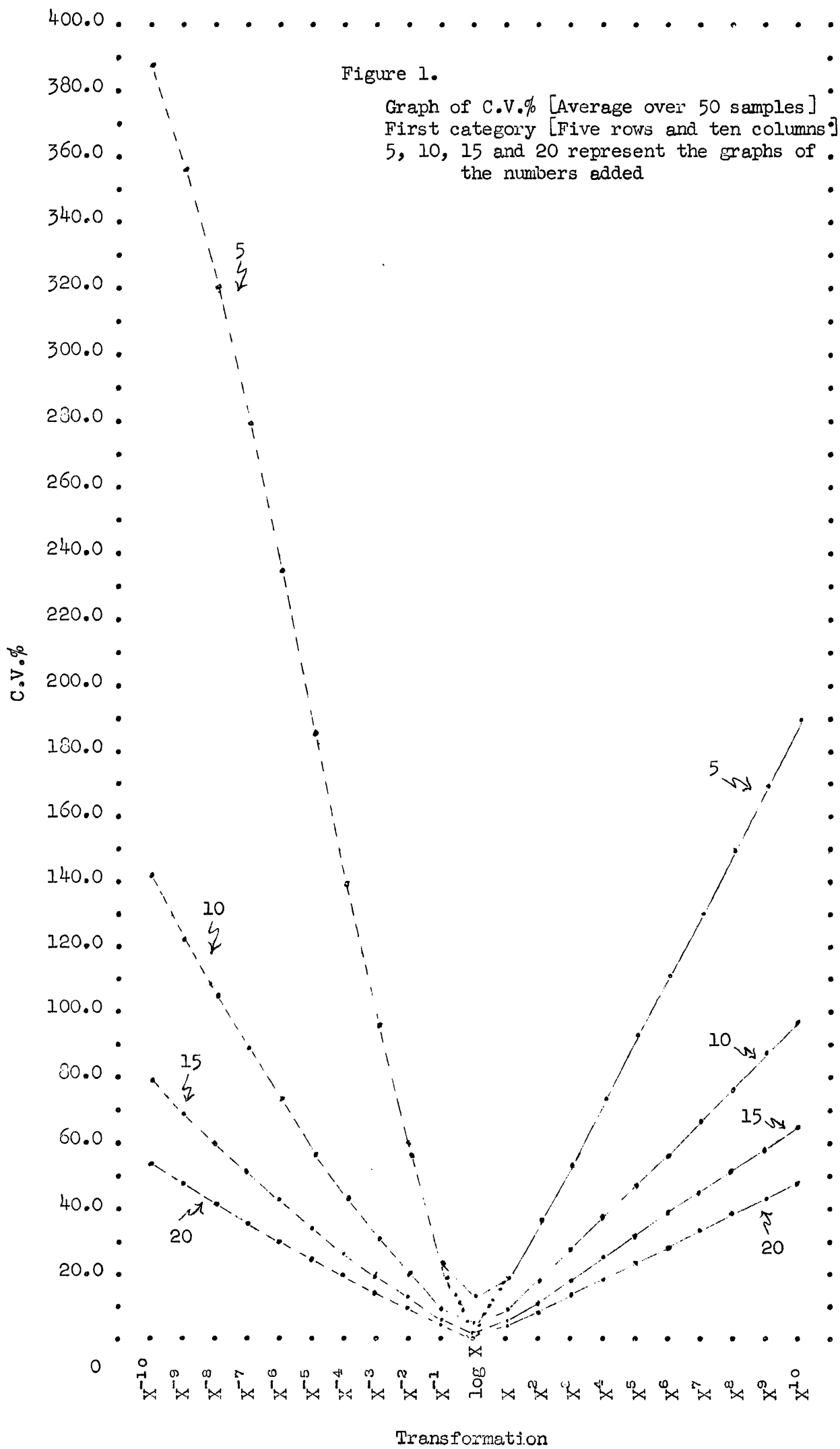
Theorem 3: Let $\log X$ be normally distributed with mean μ and variance $\sigma^2 (\sigma > 0)$. Let X^k be a transformation for all real k . Then the coefficient of variation of X^k increases as $|k|$ increases and coefficient of variation of $X^k = \text{coefficient of variation of } X^{-k}$.

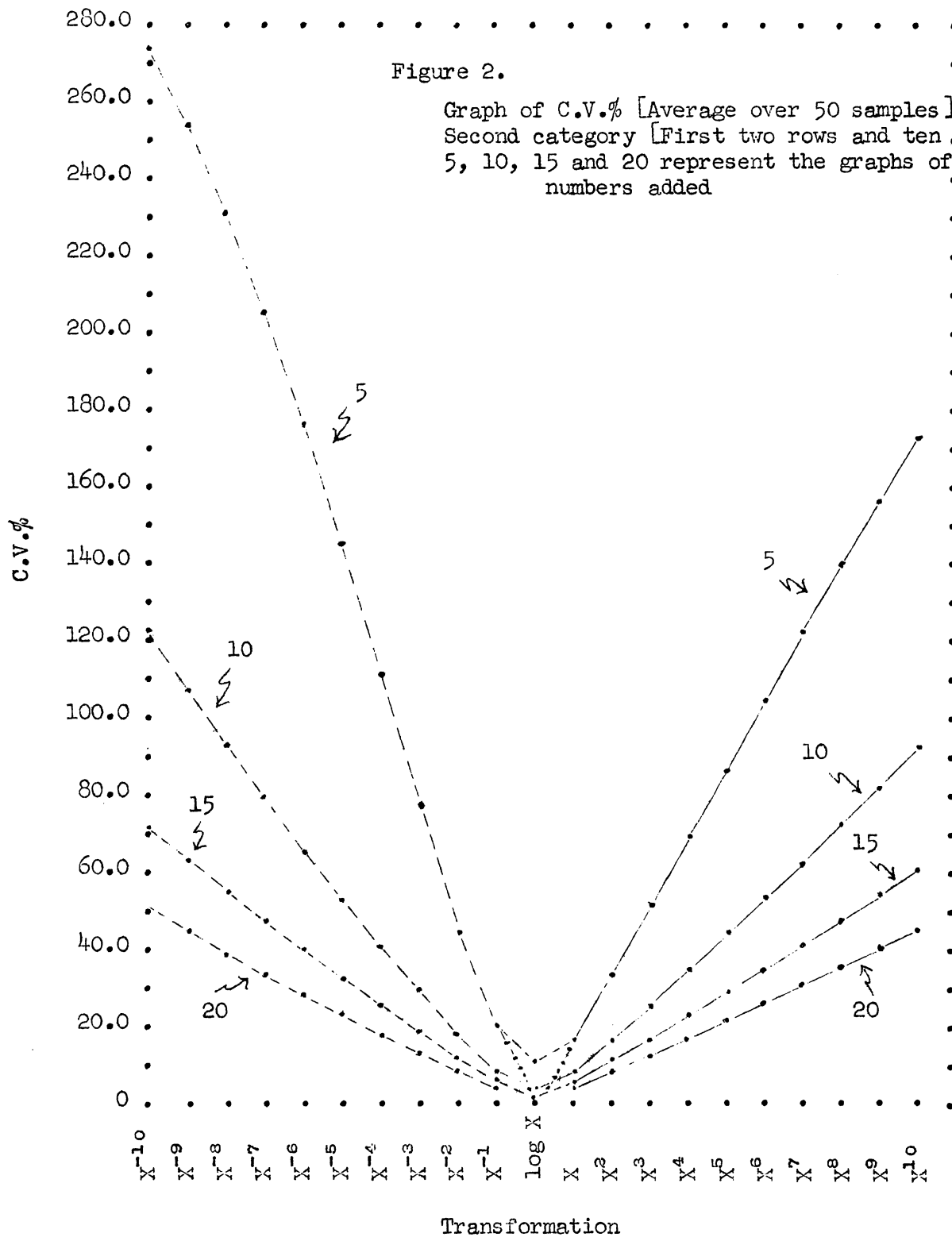
In addition an empirical study was conducted with a number of transformations of the form X^k , for $k=\pm 1, \pm 2, \dots, \pm 10$ and $k=\pm 0.2, \pm 0.4, \pm 0.5, \pm 0.6$ and ± 0.8 on a number of random samples of two-way classification drawn from normal populations. The samples were obtained in the following way. Five pages of normal deviates were selected at random from the book, "A million random digets and 100,000 normal deviates," published by The Rand Corporation. Each sample is a group of observations arranged in five rows and ten columns. To each observation in the sample numbers 5, 10, 15 and 20 were added (i.e., population mean 5, 10, 15 and 20 and variance 1 in each). The samples so obtained were designated as category 1. From each sample of category 1 the

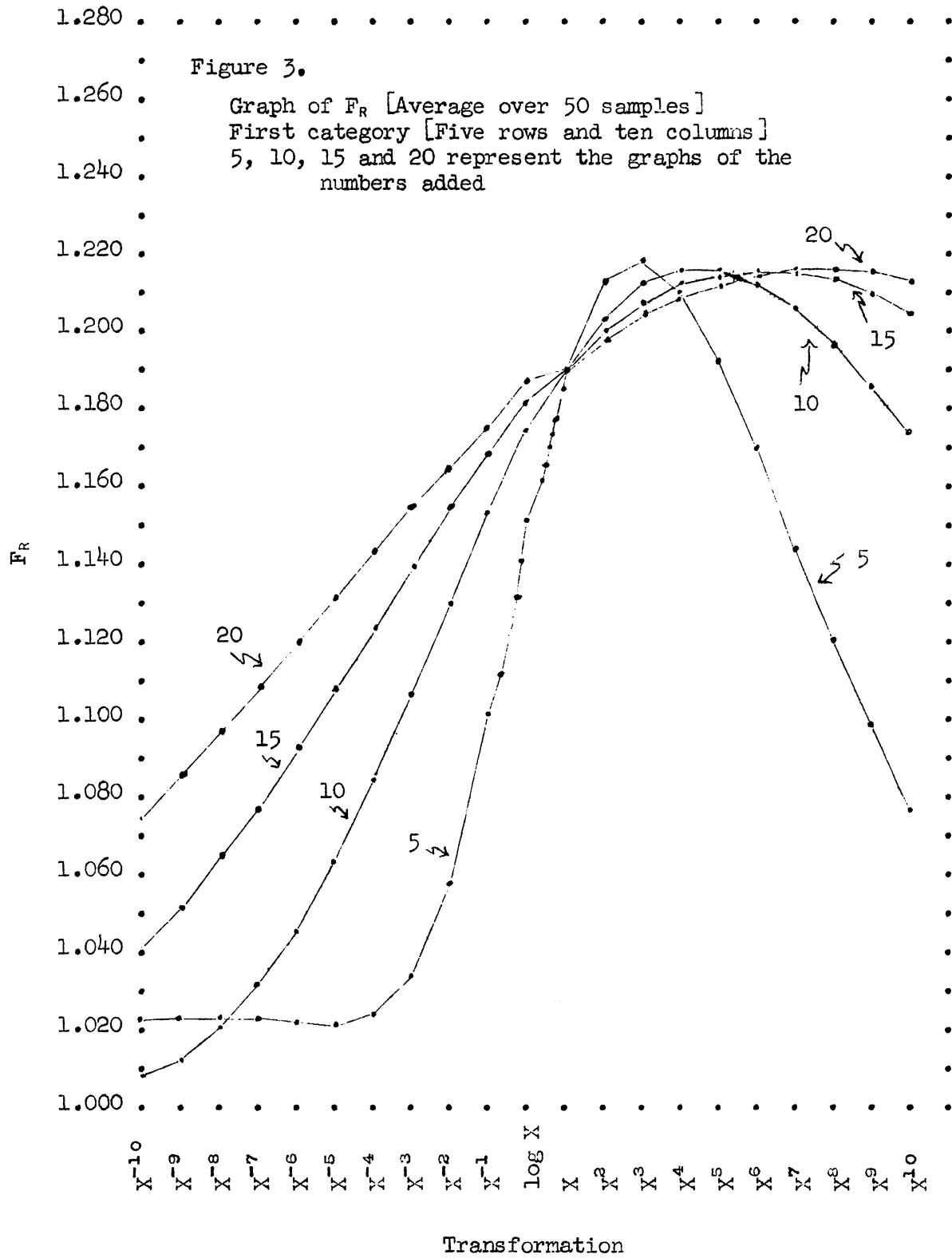
first two rows and ten columns were taken and were designated as category 2. The study of the fractional transformations was done on the samples with the population mean 5 in both categories.

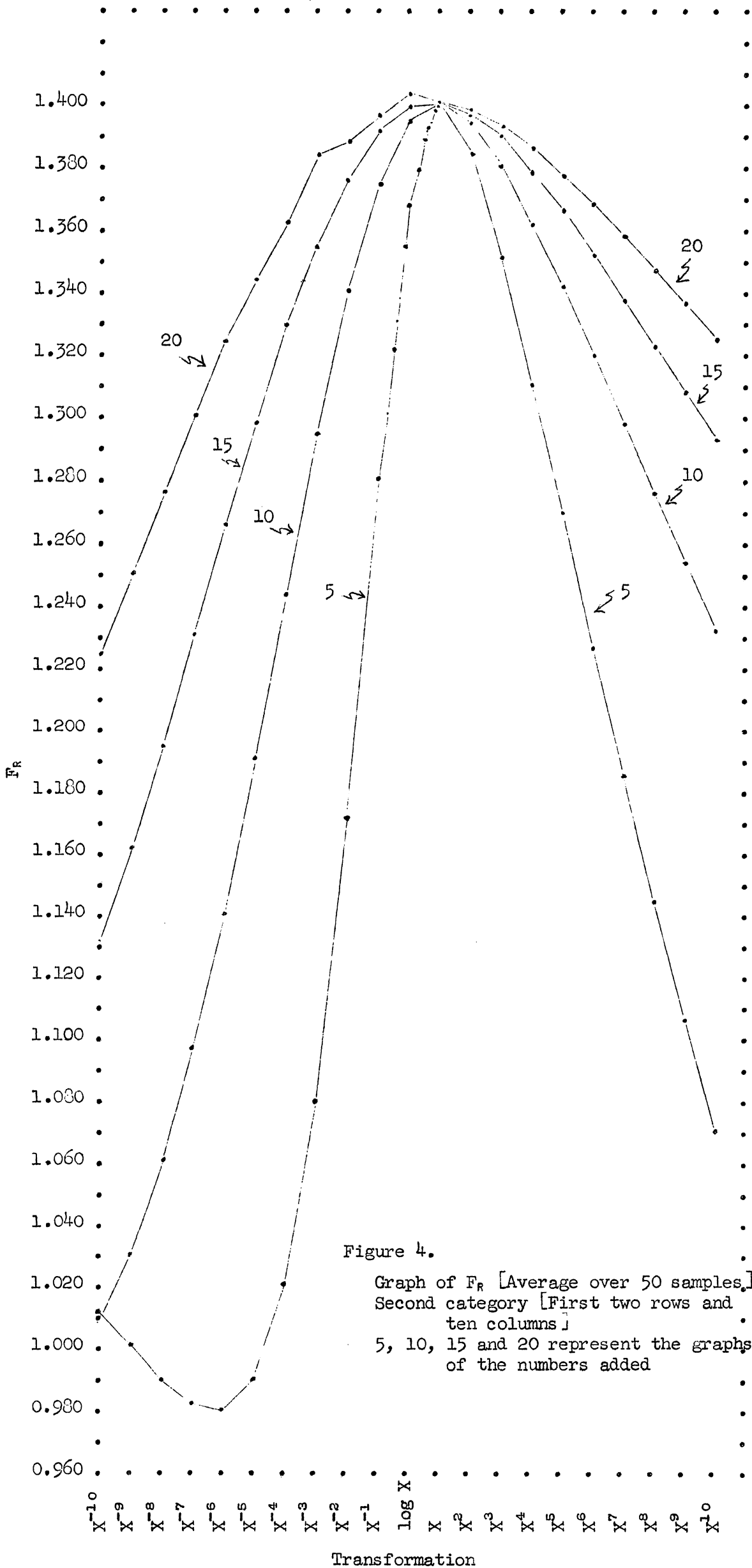
The F ratio statistics: $F_R(4,36), F_R(1,9)$ (row F); $F_C(9,36), F_C(9,9)$ (column F); $F_T(1,35), F_T(1,8)$ (Tukey's non-additivity F) and the coefficient of variation $\left(\frac{\sqrt{\text{Error mean square}}}{\text{Sample mean}} \times 100 \right)$ were calculated from the analyses of variance and were studied individually and in relation to the change in coefficient of variation.

From these results a number of conjectures and conclusions were drawn. The chief conclusion is that the minimum C.V. may be used to detect normality and to indicate the value of k in X^k which is to be analyzed.









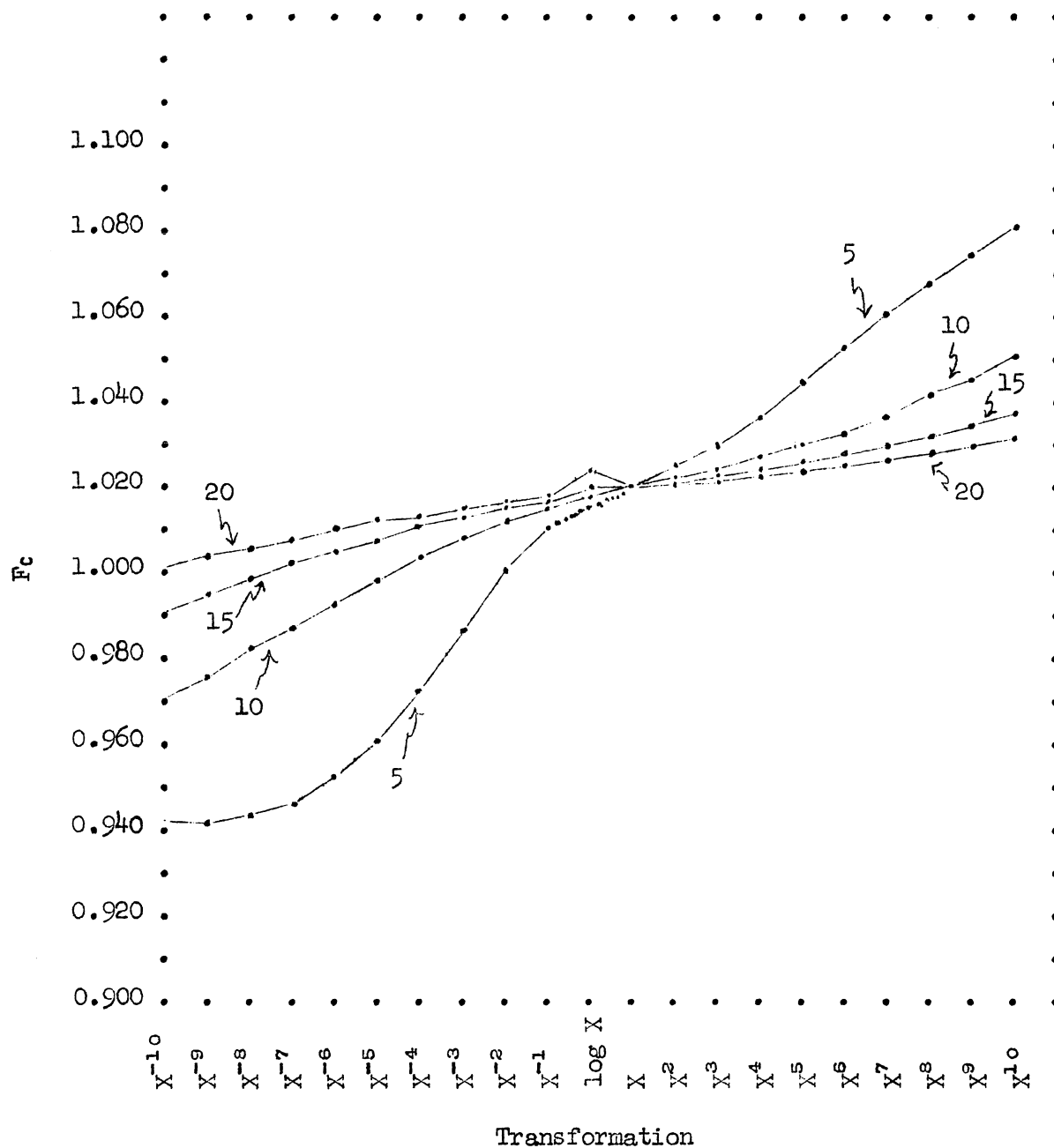
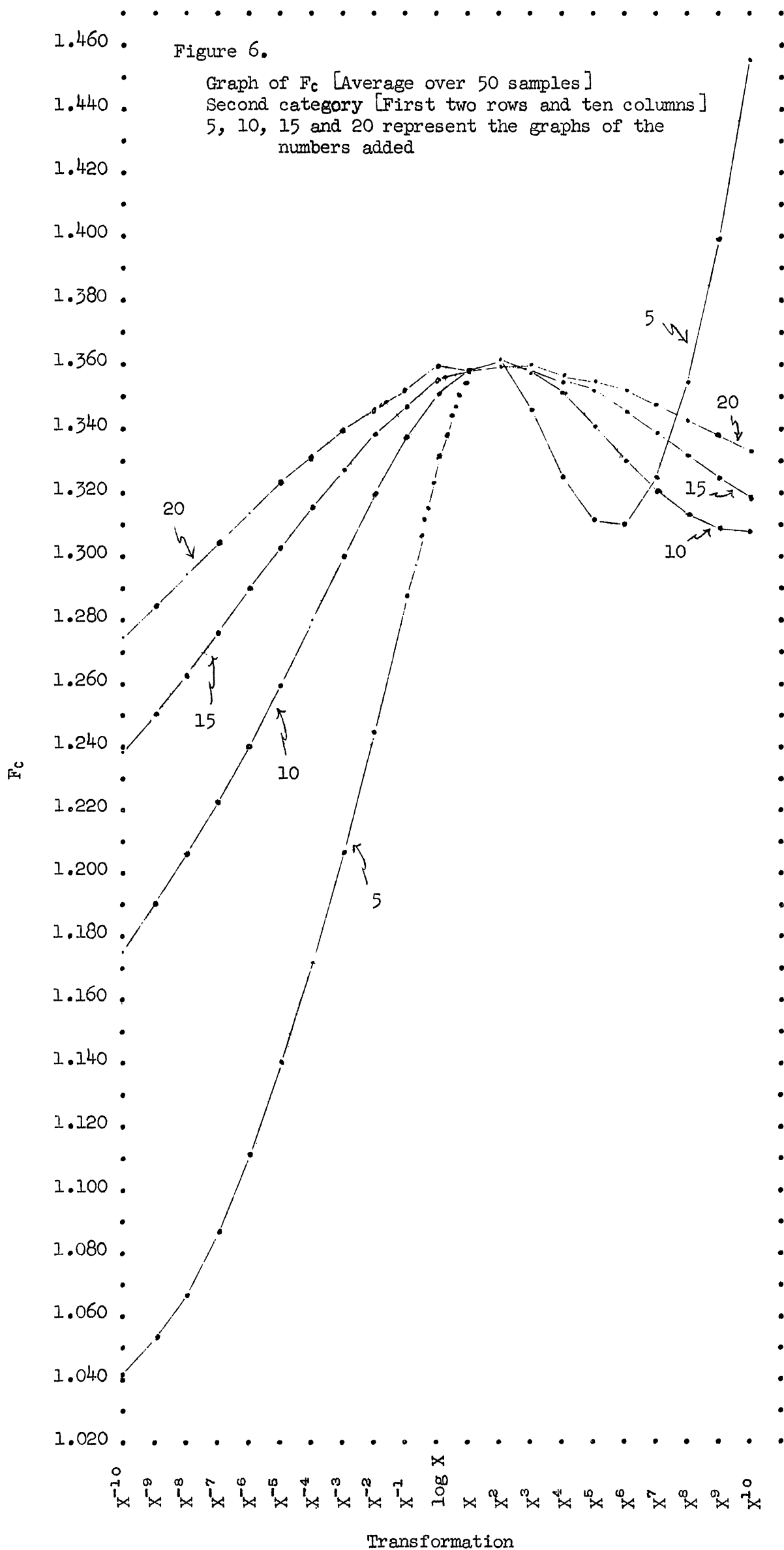


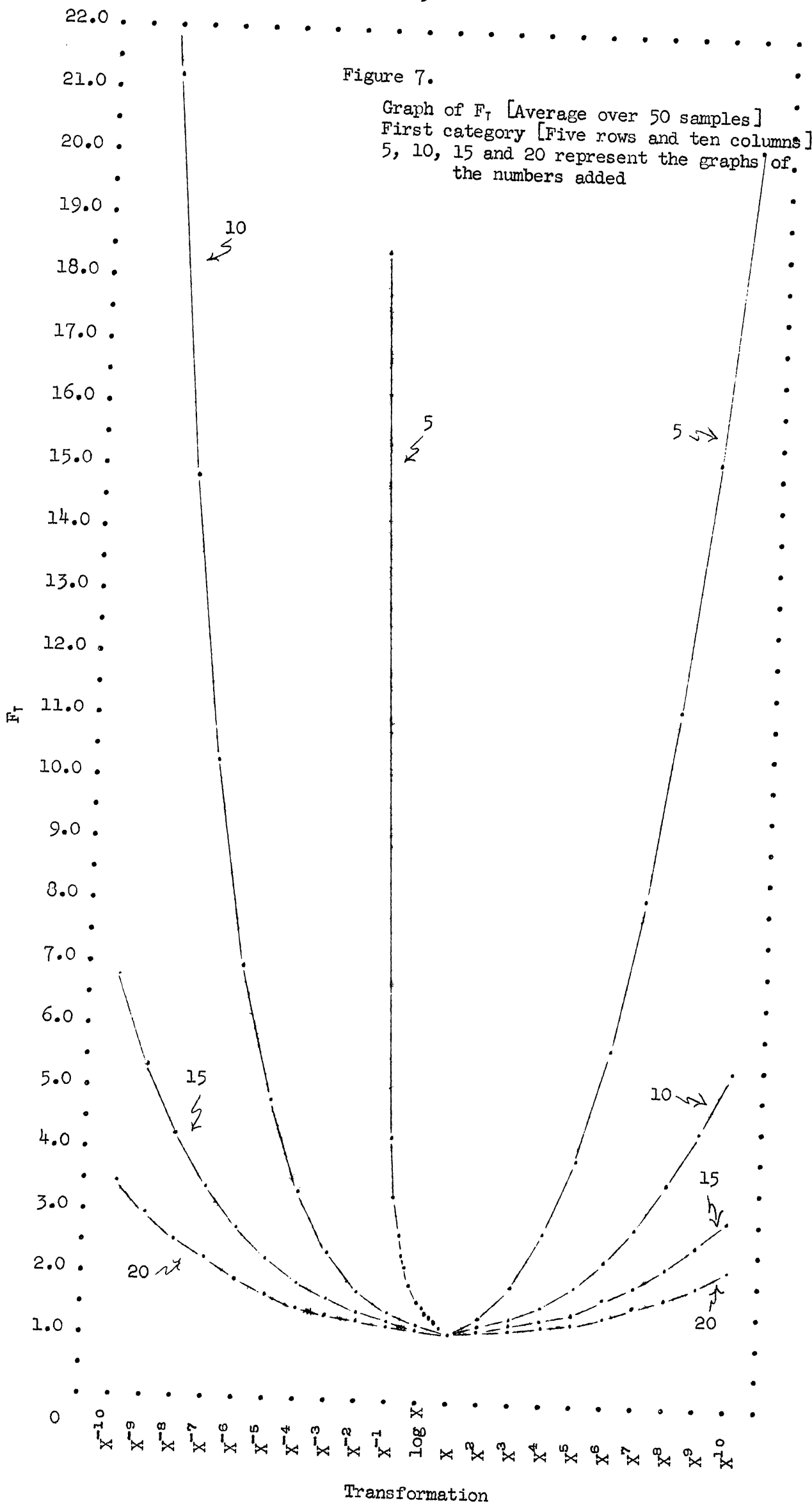
Figure 5.

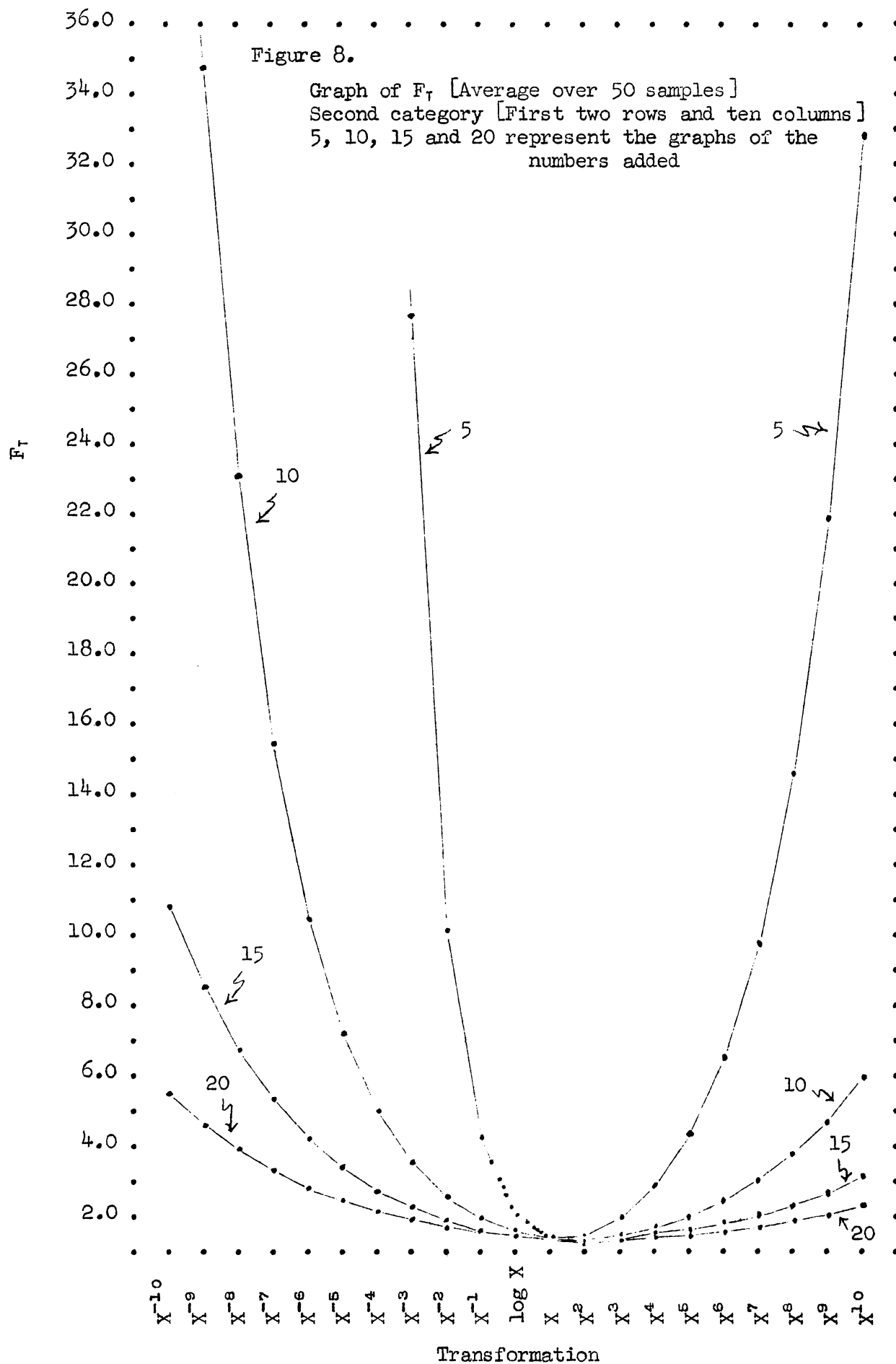
Graph of F_c [Average over 50 samples]

First category [Five rows and ten columns]

5, 10, 15 and 20 represent the graphs of the numbers added







LITERATURE CITED

- [1] Bartlett, M. S., "Square root transformation in analysis of variance," Journal of the Royal Statistical Society, Suppl. 3 (1936).
- [2] Bartlett, M. S., "The use of transformations," Biometrics, Vol. 3, No. 1, (1947).
- [3] Box, G. E. P., and Anderson, S. L., "Robust tests for variances and effects of non-normality and variance heterogeneity on standard tests," Consolidated University of North Carolina Institute of Statistics Mimeo Series No. 101 (1954).
- [4] Cochran, W. G., "Some difficulties in the statistical analysis of replicated experiments," Empire Journal of Experimental Agriculture, Vol. 6 (1938).
- [5] Cochran, W. G., "Some consequences when the assumptions for the analysis of variance are not satisfied," Biometrics, Vol. 3, No. 1 (1947).
- [6] Curtiss, J. H., "On transformations used in analysis of variance," Annals of Mathematical Statistics, 14 (1943).
- [7] Harter, H. L., and Lum, M. D., "An interpretation and extension of Tukey's one degree of freedom for non-additivity," Aeronautical Research Laboratory, Project No. 7071, Task No. 70429 (1962).
- [8] Nissen, O., and Ottestad, P., "On the analysis of variance and the effect of non-normality," Soertrykk av Meldinger fra Norges Landbrukshøgskole (1943).
- [9] Pearson, E. S., "The analysis of variance in cases of non-normal variation," Biometrika, Vol. XXIII (1931).
- [10] Tukey, J. W., "One degree of freedom for non-additivity," Biometrics, Vol. 5 (1949).
- [11] Tukey, J. W., "On the comparative anatomy of transformations," Annals of Mathematical Statistics, Vol. 28 (1957).
- [12] The Rand Corporation, "A million random digets with 100,000 normal deviates," (1955).
- [13] Rao, B. M., "Some properties of the coefficient of variation and F statistics with respect to transformations of the form X^k ," M. S. Thesis (1962), Cornell University, Ithaca, New York.

APPENDIX A

Notation: Let x be a random variable distributed normally with the mean μ and variance σ^2 ($N(\mu, \sigma^2)$) where $\mu > 0$ and $\sigma > 0$. Let k be an integer greater than or equal to 1 (i.e., $k=1,2,3,\dots$). Let α_k be the k^{th} moment of the random variable x

$$\text{(i.e., } \alpha_k = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x^k e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \text{)}$$

for $k=1,2,3,\dots$. α_0 is defined as equal to 1 and α_{-k} is defined as equal to zero.

Lemma: If x is a random variable distributed $N(\mu, \sigma^2)$ ($\sigma > 0$) then
 $\alpha_k = \mu \alpha_{k-1} + (k-1)\sigma^2 \alpha_{k-2}$

Proof:

$$\alpha_k = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x^k e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

Let $x-\mu=y$; then $x=y+\mu$ and $dx=dy$. Hence

$$\begin{aligned} \alpha_k &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (y+\mu)^k e^{-\frac{1}{2\sigma^2}y^2} dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (y+\mu)^{k-1} y e^{-\frac{1}{2\sigma^2}y^2} dy + \frac{\mu}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (y+\mu)^{k-1} e^{-\frac{1}{2\sigma^2}y^2} dy \end{aligned}$$

Integrating by parts the first part of the right hand side we get

$$\begin{aligned}
 \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (y+\mu)^{k-1} y e^{-\frac{1}{2\sigma^2}y^2} dy &= \frac{-\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (y+\mu)^{k-1} dy e^{-\frac{1}{2\sigma^2}y^2} \\
 &= \frac{-\sigma}{\sqrt{2\pi}} \left[(y+\mu)^{k-1} e^{-\frac{1}{2\sigma^2}y^2} \right]_{-\infty}^{\infty} + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (k-1)(y+\mu)^{k-2} e^{-\frac{1}{2\sigma^2}y^2} dy \\
 &= \sigma^2(k-1) \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (y+\mu)^{k-2} e^{-\frac{1}{2\sigma^2}y^2} dy \\
 &= \sigma^2(k-1)\alpha_{k-2} .
 \end{aligned}$$

The second part is $\mu\alpha_{k-1}$. Hence $\alpha_k = \mu\alpha_{k-1} + (k-1)\sigma^2\alpha_{k-2}$.

If x is a random variable then $f(x)$ is also a random variable and its expectation is

$$Ef(x) = \int f(x)dF \quad \text{where } F \text{ is the distribution function} \quad (1)$$

of the random variable x .

$$E[f(x)]^2 = \int [f(x)]^2 dF \quad (2)$$

variance of $f(x)$ is

$$E[f(x)]^2 - [Ef(x)]^2 \quad (3)$$

The standard deviation is the square root of (3).

The coefficient of variation is defined as s/\bar{x} where s is the standard deviation and \bar{x} is the mean of the observed sample. The population coefficient of variation can be written as σ/μ (i.e., standard deviation divided by mean for any $f(x)$).

Theorem: Let x be a random variable $N(\mu, \sigma^2)$, $\mu > 0$, $\sigma > 0$ and let x^k be the transformation of x for $k=1, 2, 3, \dots$. Then, the coefficient of variation of x^k increases as k increases provided $3\mu^4 > \sigma^4$.

Proof: We have to prove

$$\frac{[\alpha_{2k+2} - \alpha_{k+1}^2]^{1/2}}{\alpha_{k+1}} > \frac{[\alpha_{2k} - \alpha_k^2]^{1/2}}{\alpha_k} \quad \text{for } k=1, 2, 3, \dots$$

Because all these moments are positive since $\mu > 0$ we look at

$$\frac{\alpha_{2k+2} - \alpha_{k+1}^2}{\alpha_{k+1}^2} > \frac{\alpha_{2k} - \alpha_k^2}{\alpha_k^2}$$

Rewriting the above,

$$\frac{\alpha_{2k+2}}{\alpha_{k+1}^2} - 1 > \frac{\alpha_{2k}}{\alpha_k^2} - 1$$

and

$$\frac{\alpha_{2k+2}}{\alpha_{k+1}^2} > \frac{\alpha_{2k}}{\alpha_k^2}$$

If we can prove that $\alpha_{2k+2} \cdot \alpha_k^2 - \alpha_{2k} \cdot \alpha_{k+1}^2 > 0$, this is sufficient to prove the theorem.

From the Lemma we have

$$\alpha_{2k+2} = \mu \alpha_{2k+1} + (2k+1) \sigma^2 \alpha_{2k}$$

$$= \mu \{ \mu \alpha_{2k} + (2k) \sigma^2 \alpha_{2k-1} \} + (2k+1) \sigma^2 \alpha_{2k}$$

$$= \{ \mu^2 + (2k+1) \sigma^2 \} \alpha_{2k} + (2k) \sigma^2 \mu \alpha_{2k-1}$$

$$\alpha_{k+1} = \mu \alpha_k + k \sigma^2 \alpha_{k-1}$$

$$\alpha_{k+1}^2 = \mu^2 \alpha_k^2 + 2k \mu \sigma^2 \alpha_k \alpha_{k-1} + k^2 \sigma^4 \alpha_{k-1}^2$$

Hence

$$\alpha_{2k+2} \cdot \alpha_k^2 = [\{ \mu^2 + (2k+1) \sigma^2 \} \alpha_{2k} + (2k) \mu \sigma^2 \alpha_{2k-1}] \cdot \alpha_k^2$$

$$\alpha_{2k} \cdot \alpha_{k+1}^2 = \alpha_{2k} \cdot [\mu^2 \alpha_k^2 + 2k \mu \sigma^2 \alpha_k \alpha_{k-1} + k^2 \sigma^4 \alpha_{k-1}^2]$$

Hence

$$\begin{aligned} \alpha_{2k+2} \cdot \alpha_k^2 - \alpha_{2k} \cdot \alpha_{k+1}^2 &= \{ \mu^2 + (2k+1) \sigma^2 \} \alpha_{2k} \cdot \alpha_k^2 + (2k) \mu \sigma^2 \alpha_{2k-1} \cdot \alpha_k^2 \\ &\quad - [\mu^2 \alpha_{2k} \cdot \alpha_k^2 + 2k \mu \sigma^2 \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-1} + k^2 \sigma^4 \alpha_{2k} \cdot \alpha_{k-1}^2] \\ &= (2k+1) \sigma^2 \alpha_{2k} \cdot \alpha_k^2 + (2k) \mu \sigma^2 \alpha_{2k-1} \cdot \alpha_k^2 - 2k \mu \sigma^2 \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-1} - k^2 \sigma^4 \alpha_{2k} \cdot \alpha_{k-1}^2 \\ &= (2k+1) \sigma^2 \alpha_{2k} \cdot \alpha_k [\mu \alpha_{k-1} + (k-1) \sigma^2 \alpha_{k-2}] + (2k) \mu \sigma^2 \alpha_{2k-1} \cdot \alpha_k^2 \\ &\quad - (2k) \mu \sigma^2 \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-1} - k^2 \sigma^4 \alpha_{2k} \cdot \alpha_{k-1}^2 \\ &= \sigma^2 \mu \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-1} + (2k+1) (k-1) \sigma^4 \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-2} + (2k) \mu \sigma^2 \alpha_{2k-1} \cdot \alpha_k^2 \\ &\quad - k^2 \sigma^4 \alpha_{2k} \cdot \alpha_{k-1}^2 \end{aligned} \tag{4}$$

Now let us divide the problem into two parts (i) k is even, (ii) k is odd.

Case (i) k is even: i.e., k=2,4,6,...

$$\alpha_{k-1}^2 \leq \beta_{k-1}^2 \leq \beta_k \cdot \beta_{k-2} = \alpha_k \cdot \alpha_{k-2}$$

where β_k equal absolute k^{th} moment.

Hence substituting $\alpha_k \cdot \alpha_{k-2}$ for α_{k-1}^2 in (4), we get

$$\alpha_{2k+2} \cdot \alpha_k^2 - \alpha_{2k} \cdot \alpha_{k+1}^2 \geq \sigma^2 \mu \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-1} \\ + (k^2 - k - 1) \sigma^4 \cdot \alpha_{2k} \cdot \alpha_k \cdot \alpha_{k-2} + (2k) \mu \sigma^2 \alpha_{2k-1} \cdot \alpha_k^2 > 0$$

Note that when k is even, the condition $3\mu^4 > \sigma^4$ is not necessary.

For example, let us consider the cases when $k=2$ and $k=4$. When $k=2$ we have to show that the coefficient of variation of x^3 is greater than the coefficient of variation x^2 .

$$Ex^2 = \mu^2 + \sigma^2$$

$$Ex^3 = \mu^3 + 3\mu\sigma^2$$

$$Ex^4 = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4$$

$$Ex^6 = 15\sigma^6 + 45\mu^2\sigma^4 + 15\mu^4\sigma^2 + \mu^6$$

We have to show that

$$\frac{Ex^6}{(Ex^3)^2} > \frac{Ex^4}{(Ex^2)^2} \quad (5)$$

That is,

$$\frac{(15\sigma^6 + 45\mu^2\sigma^4 + 15\mu^4\sigma^2 + \mu^6)}{(\mu^3 + 3\mu\sigma^2)^2} > \frac{(3\sigma^4 + 6\mu^2\sigma^2 + \mu^4)}{(\mu^2 + \sigma^2)^2}$$

or

$$(15\sigma^6 + 45\mu^2\sigma^4 + 15\mu^4\sigma^2 + \mu^6)(\mu^4 + 2\mu^2\sigma^2 + \sigma^4) - (3\sigma^4 + 6\mu^2\sigma^2 + \mu^4)(\mu^6 + 6\mu^4\sigma^2 + 9\mu^2\sigma^4) > 0$$

Expanding we obtain,

$$120\mu^4\sigma^6 + 76\mu^6\sigma^4 + 17\mu^8\sigma^2 + \mu^{10} + 75\mu^2\sigma^8 + 15\sigma^{10} - 72\mu^4\sigma^6 - 48\mu^6\sigma^4 - 12\mu^8\sigma^2 - \mu^{10} \\ - 27\mu^2\sigma^8 > 0$$

or,

$$15\sigma^{10} + 48\sigma^8\mu^2 + 48\sigma^6\mu^4 + 28\sigma^4\mu^6 + 5\sigma^2\mu^8 > 0$$

which is obviously true.

When $k=4$ we have to show that the coefficient of variation of x^5 is greater than coefficient of variation of x^4 . That is, we have to show that

$$\frac{Ex^{10}}{(Ex^5)^2} > \frac{Ex^8}{(Ex^4)^2} \quad (6)$$

$$Ex^4 = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4$$

$$Ex^5 = 15\mu\sigma^4 + 10\mu^3\sigma^2 + \mu^5$$

$$Ex^8 = 105\sigma^8 + 420\mu^2\sigma^6 + 210\mu^4\sigma^4 + 28\mu^6\sigma^2 + \mu^8$$

$$Ex^{10} = 945\sigma^{10} + 4725\mu^2\sigma^8 + 3150\mu^4\sigma^6 + 630\mu^6\sigma^4 + 45\mu^8\sigma^2 + \mu^{10}$$

Then (6) becomes

$$\begin{aligned} & \frac{(945\sigma^{10} + 4725\mu^2\sigma^8 + 3150\mu^4\sigma^6 + 630\mu^6\sigma^4 + 45\mu^8\sigma^2 + \mu^{10})}{(15\mu\sigma^4 + 10\mu^3\sigma^2 + \mu^5)^2} \\ & > \frac{(105\sigma^8 + 420\mu^2\sigma^6 + 210\mu^4\sigma^4 + 28\mu^6\sigma^2 + \mu^8)}{(3\sigma^4 + 6\mu^2\sigma^2 + \mu^4)^2} \end{aligned}$$

or,

$$\begin{aligned} & (945\sigma^{10} + 4725\mu^2\sigma^8 + 3150\mu^4\sigma^6 + 630\mu^6\sigma^4 + 45\mu^8\sigma^2 + \mu^{10}) \\ & \cdot (9\sigma^8 + 36\mu^2\sigma^6 + 42\mu^4\sigma^4 + 12\mu^6\sigma^2 + \mu^8) \\ & > (105\sigma^8 + 420\mu^2\sigma^6 + 210\mu^4\sigma^4 + 28\mu^6\sigma^2 + \mu^8)(225\mu^2\sigma^8 + 300\mu^4\sigma^6 + 130\mu^6\sigma^4 \\ & \quad + 20\mu^8\sigma^2 + \mu^{10}) \end{aligned}$$

and,

$$\begin{aligned}
 & 8505\sigma^{18} + 76545\mu^2\sigma^{16} + 238140\mu^4\sigma^{14} + 328860\mu^6\sigma^{12} + 213030\mu^8\sigma^{10} + 70614\mu^{10}\sigma^8 \\
 & + 12636\mu^{12}\sigma^6 + 1212\mu^{14}\sigma^4 + 57\mu^{16}\sigma^2 + \mu^{18} - 23625\mu^2\sigma^{16} - 126000\mu^4\sigma^{14} \\
 & - 186900\mu^6\sigma^{12} - 126000\mu^8\sigma^{10} - 44430\mu^{10}\sigma^8 - 8560\mu^{12}\sigma^6 - 900\mu^{14}\sigma^4 \\
 & - 48\mu^{16}\sigma^2 - \mu^{18} > 0
 \end{aligned}$$

Rewriting we obtain,

$$\begin{aligned}
 & 8505\sigma^{18} + 52920\mu^2\sigma^{16} + 112140\mu^4\sigma^{14} + 141960\mu^6\sigma^{12} + 87030\mu^8\sigma^{10} + 26184\mu^{10}\sigma^8 \\
 & + 4076\mu^{12}\sigma^6 + 312\mu^{14}\sigma^4 + 9\mu^{16}\sigma^2 > 0
 \end{aligned}$$

which is obviously true.

Case (ii): k is odd i.e., $k=1,3,5,\dots$

All the moments can be expressed in terms of μ and σ . Thus, we may write

$$\alpha_{2k+2} = \sum_{r=0}^{k+1} \binom{2k+2}{2r} \mu^{2r} \sigma^{2k+2-2r} \cdot 1 \cdot 3 \cdot 5 \cdots (2k+1-2r)$$

$$\alpha_{2k} = \sum_{r=0}^k \binom{2k}{2r} \mu^{2r} \sigma^{2k-2r} \cdot 1 \cdot 3 \cdot 5 \cdots (2k-2r-1)$$

$$= \sum_{r=0}^{k+1} \binom{2k}{2r-2} \mu^{2r-2} \sigma^{2k+2-2r} \cdot 1 \cdot 3 \cdot 5 \cdots (2k+1-2r)$$

$$\alpha_{k+1} = \sum_{j=0}^{\frac{k+1}{2}} \binom{k+1}{2j} \mu^{2j} \sigma^{k+1-2j} \cdot 1 \cdot 3 \cdot 5 \cdots (k-2j)$$

$$\begin{aligned}
 \alpha_{k+1}^2 &= \sum_{j=0}^{\frac{k+1}{2}} \sum_{h=0}^{\frac{k+1}{2}} \binom{k+1}{2j} \binom{k+1}{2h} \mu^{2j+2h} \sigma^{2k+2-2j-2h} \cdot 1 \cdot 3 \cdot 5 \cdots (k-2j) \\
 &\quad \times 1 \cdot 3 \cdot 5 \cdots (k-2h)
 \end{aligned}$$

$$\alpha_k = \sum_{j=1}^{\frac{k+1}{2}} \binom{k}{2j-1} \mu^{2j-1} \cdot \sigma^{k+1-2j} \cdot 1 \cdot 3 \cdot 5 \cdots (k-2j)$$

$$= \sum_{j=0}^{\frac{k+1}{2}} \binom{k}{2j-1} \mu^{2j-1} \cdot \sigma^{k+1-2j} \cdot 1 \cdot 3 \cdot 5 \cdots (k-2j)$$

$$\alpha_k^2 = \sum_{j=0}^{\frac{k+1}{2}} \sum_{h=0}^{\frac{k+1}{2}} \binom{k}{2j-1} \binom{k}{2h-1} \mu^{2j+2h-2} \cdot \sigma^{2k+2-2j-2h} \cdot 1 \cdot 3 \cdot 5 \cdots (k-2j) \\ \times 1 \cdot 3 \cdot 5 \cdots (k-2h)$$

Note: When r, j and h take their maximum limit the coefficient $1 \cdot 3 \cdot 5 \cdots (\dots)$ is to be taken as 1 and not as -1 since the power of σ becomes zero the coefficient exists as 1 only. But when $0 \leq r < k+1$, $0 \leq j < \frac{k+1}{2}$, $0 \leq h < \frac{k+1}{2}$, the coefficient $1 \cdot 3 \cdot 5 \cdots (\dots)$ is to be taken as usual.

$$\alpha_{2k+2} \cdot \alpha_k^2 - \alpha_{2k} \cdot \alpha_{k+1}^2 \\ = \sum_{r,j,h} \binom{2k+2}{2r} \binom{k}{2j-1} \binom{k}{2h-1} \mu^{2j+2h+2r-2} \cdot \sigma^{4k+4-2r-2j-2h} \\ 1 \cdot 3 \cdot 5 \cdots (2k+1-2r) \times 1 \cdot 3 \cdot 5 \cdots (k-2j) \times 1 \cdot 3 \cdot 5 \cdots (k-2h) \\ - \sum_{r,j,h} \binom{2k}{2r-2} \binom{k+1}{2j} \binom{k+1}{2h} \mu^{2j+2h+2r-2} \cdot \sigma^{4k+4-2j-2h-2r} \\ 1 \cdot 3 \cdot 5 \cdots (2k+1-2r) \times 1 \cdot 3 \cdot 5 \cdots (k-2j) \times 1 \cdot 3 \cdot 5 \cdots (k-2h) \\ = \sum_{r,j,h} 1 \cdot 3 \cdot 5 \cdots (2k+1-2r) \times 1 \cdot 3 \cdot 5 \cdots (k-2j) \times 1 \cdot 3 \cdot 5 \cdots (k-2h) \mu^{2j+2h+2r-2} \\ \cdot \sigma^{4k+4-2r-2j-2h} \left[\binom{2k+2}{2r} \binom{k}{2j-1} \binom{k}{2h-1} - \binom{2k}{2r-2} \binom{k+1}{2j} \binom{k+1}{2h} \right] \quad (7)$$

In $\alpha_{2k} \cdot \alpha_{k+1}^2$ there is a term $\sigma^{4k+2} 1 \cdot 3 \cdot 5 \cdots (k) \times 1 \cdot 3 \cdot 5 \cdots (k) \times 1 \cdot 3 \cdot 5 \cdots (2k-1)$ which is not in $\alpha_{2k+2} \cdot \alpha_k^2$. Hence consider the terms involving μ and σ in the above summation. The first thing that is required to be proved is that the terms involving μ and σ are positive in (7). After this a condition can be

imposed which makes (7) definitely positive.

In order to prove that the terms involving μ and σ in (7) are positive it is sufficient to consider the coefficient

$$\binom{2k+2}{2r} \binom{k}{2j-1} \binom{k}{2h-1} - \binom{2k}{2r-2} \binom{k+1}{2j} \binom{k+1}{2h} \quad (8)$$

and show that it is positive. After some simplification (8) turns out to be

$$(2k+1)4jh - (k+1)r(2r-1) \quad (9)$$

Treating (9) as quadratic in r it is required to show that

$$2(k+1)r^2 - (k+1)r - (2k+1)4jh \leq 0 \quad (10)$$

For (10) to be negative, the discriminant should be positive which is true and r should lie in the interval $[0, k+1]$ (i.e., $0 \leq r \leq k+1$). Because these are the limits of summation, r does lie between 0 and $k+1$. Hence negativeness of (10) implies positiveness of (9).

In (7) the terms $\mu^2 \sigma^{4k}$ and μ^{4k+2} vanish. From (7) the term $\mu^4 \sigma^{4k-2}$ has the coefficient $\frac{(7k^2+8k+3)}{6} 1 \cdot 3 \cdot 5 \cdots (2k-1) 1 \cdot 3 \cdot 5 \cdots k \times 1 \cdot 3 \cdot 5 \cdots k$. When $k=1$ terms containing $\mu^4 \sigma^2$ and σ^6 only remain. Hence it is proper to consider these two terms only in order to have a condition to make (7) strictly positive. Considering the terms $\frac{(7k^2+8k+3)}{6} 1 \cdot 3 \cdot 5 \cdots (2k-1) 1 \cdot 3 \cdot 5 \cdots k \times 1 \cdot 3 \cdot 5 \cdots k \mu^4 \sigma^{4k-2}$ and $1 \cdot 3 \cdot 5 \cdots (2k-1) 1 \cdot 3 \cdot 5 \cdots (k) \times 1 \cdot 3 \cdot 5 \cdots k \sigma^{4k+2}$ we get the condition that $\frac{(7k^2+8k+3)}{6} \mu^4 > \sigma^4$. When $k=1$ this condition turns out to be $3\mu^4 > \sigma^4$ which is the condition of the theorem. Rewriting the inequality as $\mu^4 > \frac{6}{7k^2+8k+3} \sigma^4$, we note that for sufficiently large k the inequality becomes essentially $\mu > 0$.

To illustrate the above, two examples are given for $k=1$ and $k=3$.

Case (i) $k=1$

$$\begin{aligned} Ex = \mu & \quad ; \quad Ex^2 = \mu^2 + \sigma^2 \\ & \quad ; \quad Ex^4 = 3\sigma^4 + 6\mu^2 \sigma^2 + \mu^4 \end{aligned}$$

With the condition $3\mu^4 > \sigma^4$ we will show that

$$\frac{Ex^4}{[Ex^2]^2} > \frac{Ex^2}{[Ex]^2} \quad (11)$$

Substituting the respective values in (11) we get

$$\frac{3\sigma^4 + 6\mu^2\sigma^2 + \mu^4}{(\mu^2 + \sigma^2)^2} > \frac{(\mu^2 + \sigma^2)}{\mu^2}$$

$$3\mu^2\sigma^4 + 6\mu^4\sigma^2 + \mu^6 - \mu^6 - 3\mu^4\sigma^2 - 3\mu^2\sigma^4 - \sigma^6 > 0$$

$$3\mu^4\sigma^2 - \sigma^6 > 0$$

$$\sigma^2(3\mu^4 - \sigma^4) > 0$$

which is true because of the condition $3\mu^4 > \sigma^4$.

Case (ii) When $k=3$

$$Ex^3 = \mu^3 + 3\mu\sigma^2, \quad Ex^6 = 15\sigma^6 + 45\mu^2\sigma^4 + 15\mu^4\sigma^2 + \mu^6$$

$$Ex^4 = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4, \quad Ex^8 = 105\sigma^8 + 420\mu^2\sigma^6 + 210\mu^4\sigma^4 + 28\mu^6\sigma^2 + \mu^8$$

$$\frac{Ex^8}{[Ex^4]^2} > \frac{Ex^6}{[Ex^3]^2} \quad (12)$$

Substituting the values in (12) we get

$$(105\sigma^8 + 420\mu^2\sigma^6 + 210\mu^4\sigma^4 + 28\mu^6\sigma^2 + \mu^8)(\mu^6 + 6\mu^4\sigma^2 + 9\mu^2\sigma^4)$$

$$- (15\sigma^6 + 45\mu^2\sigma^4 + 15\mu^4\sigma^2 + \mu^6)(9\sigma^8 + 36\mu^2\sigma^6 + 42\mu^4\sigma^4 + 12\mu^6\sigma^2 + \mu^8) > 0$$

$$= 2025\mu^4\sigma^{10} + 1896\mu^6\sigma^8 + 711\mu^8\sigma^6 + 120\mu^{10}\sigma^4 + 7\mu^{12}\sigma^2 - 135\sigma^{14} > 0$$

which can easily be seen true with the condition that $3\mu^4 > \sigma^4$.

Note: As mentioned above, there is a term containing only a power of σ which becomes negative and the rest of the terms in (7) are positive. Also, the

terms involving μ^2 and a power of σ and terms involving only a power of μ vanish. This can be observed in the above two examples.

APPENDIX B

Theorem 2: Let X be a random variable normally distributed with mean μ and variance σ^2 ($\sigma > 0$). Let α_k be the k^{th} moment. Then $\alpha_{2k} \leq \alpha_k^2 \left[1 + k^2 \frac{\sigma^2}{\mu^2} \right]$ for $k=1,2,3,4$ and 5 with equality when $k=1$ only.

Proof: When $k=1$

$$\alpha_2 = \alpha_1^2 \left[1 + \frac{\sigma^2}{\mu^2} \right] = \mu^2 + \sigma^2$$

When $k=2$

$$\alpha_4 < \alpha_2^2 \left[1 + 4 \frac{\sigma^2}{\mu^2} \right]$$

$$\alpha_4 = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4$$

$$\alpha_2^2 \left[1 + 4 \frac{\sigma^2}{\mu^2} \right] = 9\sigma^4 + 6\mu^2\sigma^2 + \mu^4 + 4\sigma^6/\mu^2$$

When $k=3$

$$\alpha_6 < \alpha_3^2 \left[1 + 9 \frac{\sigma^2}{\mu^2} \right]$$

$$\alpha_6 = 15\sigma^6 + 45\sigma^4\mu^2 + 15\sigma^2\mu^4 + \mu^6$$

$$\alpha_3^2 \left[1 + 9 \frac{\sigma^2}{\mu^2} \right] = 81\sigma^6 + 63\sigma^4\mu^2 + 15\sigma^2\mu^4 + \mu^6$$

When $k=4$

$$\alpha_8 < \alpha_4^2 \left[1 + 16 \frac{\sigma^2}{\mu^2} \right]$$

$$\alpha_8 = 105\sigma^8 + 420\sigma^6\mu^2 + 210\sigma^4\mu^4 + 28\sigma^2\mu^6 + \mu^8$$

$$\alpha_4^2 \left[1 + 16 \frac{\sigma^2}{\mu^2} \right] = 585\sigma^8 + 708\sigma^6\mu^2 + 234\sigma^4\mu^4 + 28\sigma^2\mu^6 + \mu^8 + 144\sigma^{10}/\mu^2$$

When $k=5$

$$\alpha_{10} < \alpha_5^2 \left[1 + 25 \frac{\sigma^2}{\mu^2} \right]$$

$$\alpha_{10} = 945\sigma^{10} + 4725\sigma^8\mu^2 + 3150\sigma^6\mu^4 + 630\sigma^4\mu^6 + 45\sigma^2\mu^8 + \mu^{10}$$

$$\alpha_5^2 \left[1 + 25 \frac{\sigma^2}{\mu^2} \right] = 5625\sigma^{10} + 7725\sigma^8\mu^2 + 3550\sigma^6\mu^4 + 630\sigma^4\mu^6 + 45\sigma^2\mu^8 + \mu^{10}$$

By comparing the respective coefficients **one** can obviously see the truth of the theorem.

APPENDIX C

Theorem 3: Let $\log X$ be normally distributed with mean μ and variance σ^2 ($\sigma > 0$). Let X^k be a transformation for all real k . Then the coefficient of variation of X^k increases as $|k|$ increases and coefficient of variation of $X^k =$ coefficient of variation of X^{-k} .

Proof:

$$X = e^{\log X}$$

$$X^k = e^{k \log X} ; EX^k = Ee^{k \log X}$$

$$\begin{aligned} EX^k &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{k \log x} e^{-\frac{1}{2\sigma^2} (\log x - \mu)^2} dx \\ &= e^{k\mu + \frac{1}{2} k^2 \sigma^2} \end{aligned}$$

$$\text{Coefficient of Variation of } X^k = \frac{(e^{2k\mu + 2k^2\sigma^2} - e^{2k\mu + k^2\sigma^2})^{1/2}}{e^{k\mu + \frac{1}{2} k^2 \sigma^2}}$$

$$= (e^{k^2\sigma^2} - 1)^{1/2} \quad [\text{Positive root only is considered}]$$

The assertion of the theorem follows immediately.

Note: The coefficient of variation is independent of μ .